For Reference

NOT TO BE TAKEN FROM THIS ROOM

For Reference

NOT TO BE TAKEN FROM THIS ROOM

Ex LIBRIS
UNIVERSITATIS
ALBERTAENSIS

UNIVERSITY OF ALBERTA
QUAECUMQUE VERA

University of Alberta
Printing Department

Thesis
1963(F)
#70

THE UNIVERSITY OF ALBERTA

SYSTEMS ANALYSIS IN EDUCATION: A FACTOR ANALYSIS AND ANALYSIS OF VARIANCE OF PUPIL ACHIEVEMENT

by

THOMAS BARR GREENFIELD

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE

OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF EDUCATIONAL ADMINISTRATION

EDMONTON, ALBERTA

AUGUST, 1963

ABSTRACT

General systems theory provides a framework for the analysis of large organizations. In this theory organized complexities are regarded as wholes whose various parts interact to produce some output. An important consequence of systems theory suggests that a system maintains output at constant values despite varying inputs into the system and varying conditions external to the system. Differences in output associated with levels of a system may be regarded as due to system effects.

The organization of education may be defined as a system. One purpose of the study was to assign the effects which various levels of organization within an educational system have upon output. The output selected was the achievement of pupils as measured by academic examinations. The design of the study involved an hierarchical arrangement of classes, schools, and districts. The sample was made up of twenty-two school districts. In each district, two schools were selected as part of the sample, and, in each school, two classes were selected. In this way, the sample was formed of the twenty-two districts, forty-four schools, and eighty-eight classrooms. These classes, schools, and districts were regarded as sub-systems of the total system. The sample included 2,069 Grade IX pupils.

An analysis of variance design permitted total variation in achievement to be broken into sources from classes, schools, districts and a residual source from pupils within classes. In the analysis of variance on sets of achievement data from five academic subjects,

Digitized by the Internet Archive
in 2018 with funding from
University of Alberta Libraries

https://archive.org/details/systemsanalysisi00thom

ACKNOWLEDGEMENTS

The author is indebted to members of his committee for their direction, criticism, and encouragement of this thesis. The work of Dr. J. H. M. Andrews and Dr. J. A. McGregor is acknowledged for their help in problems of design and analysis. Dr. S. M. Hunka and Mrs. P. Montgomery were of assistance in problems of factor analysis and computer programming.

Thanks are also due for their generous co-operation to the Alberta Department of Education and to the staff and students of many schools.

Financial assistance for the study from the Canadian Education Association and the Department of Educational Administration is gratefully acknowledged. Fellowships from the Province of Alberta and the University of Alberta provided assistance for graduate study at the University of Alberta.

The author is also indebted to his wife for her assistance in the collection of data and the preparation of the report of this thesis.

TABLE OF CONTENTS

CHAPTER		PAGE
I.	THE PROBLEM	1
	Systems in Education	3
	Difficulties in Applying Systems Analysis to Education	7
	An Example of a System in Education	10
	The Purposes of the Study	11
	Definitions	12
	System	12
	Sub-system	12
	Elements	12
	Relationships	12
	Design of the Study	12
	Analysis of Variance	13
	Factor Analysis	13
	The Intercorrelation Matrix	14
	Hypotheses	14
II.	THE NATURE OF SYSTEMS ANALYSIS	16
	General Systems	16
	Consequences of Systems Theory	17
	Characteristics of Systems	21
	Objects	21
	Attributes	21
	Relationships	21
	Boundary	22

CHAPTER PAGE

Environment 22
Sub-system 23
Open and Closed Systems 23
Steady state 24
Equifinality 25
Feedback 27
Wholeness and independence 29
Optimization 30
Systems and Organization 30
The Iron Laws of Organization 34
The Malthusian law of growth 34
The law of culture patterns and economic activity . . . 35
The law of optimum size 35
The law of hierarchy 35
The law of oligopoly 36
The law of instability in uncontrolled interaction . . . 36
The law of role persistence 37
Communication and System 37
Failure in receptors 37
Failure in transmission 38
Failure in the executive 38
Applications of Systems Analysis in Education 39
Summary 43

CHAPTER PAGE

III. THEORETICAL FRAMEWORK FOR THE ANALYSIS OF DATA 45

Analysis of Variance 45

Image Analysis 47

The Derivation of Factor Scores 52

Scalogram Analysis 54

The Measurement of Socio-Economic Status 57

Occupational scales 59

The Home Index Scale 59

The Measurement of Educational Ability 60

IV. THE COLLECTION OF DATA 63

Data from Questionnaires 64

Achievement Data 64

The Sample 65

V. THE TEST DATA 67

Tests Used in the Study 67

The Home Index Scale 67

The Blishen Occupational Scale 68

The Socio-Economic Scale 69

Study hours 69

Motivation 70

The Alberta Departmental Examinations 70

The Sequential Tests of Educational Progress: Reading. . 72

The Cooperative School and College Ability Tests 73

Summary and enumeration of tests 74

The Use of Computer Programs 75

CHAPTER PAGE

Intercorrelation program 75

Inverse of matrix program 75

Eigenvalue-eigenvector program 75

Varimax rotation program 76

Standard score program 76

Matrix multiplication program 77

Data conversion program 77

Analysis of variance program 77

Card Punching and Handling 78

VI. THE ANALYSIS OF DATA 80

Analysis of Variance 80

Assumptions of analysis of variance 81

Image Analysis 85

Principal axes factoring 86

The meaning of principal axes factors 87

Rotation of Axes 87

The Derivation of Factor Scores 88

The analysis of factor scores 90

Scalogram Analysis 90

Rapid machine methods of scalogram analysis 94

The scale items. 94

The Reporting of Findings 97

VII. FINDINGS RELATED TO IMAGE ANALYSIS AND PRINCIPAL AXES FACTORING 98

Image Analysis 98

CHAPTER PAGE

The intercorrelation matrix 99

The inverse of the intercorrelation matrix 99

The anti-image matrix 103

The image covariance matrix 105

Principal Axes Factoring 107

Eigenvalues and eigenvectors 107

The factor matrix 109

The Rotation of Axes 111

The varimax rotation 113

Interpretation of Factors 113

Discussion 116

Test selection 118

The Hypotheses 119

Hypothesis 4 120

Hypothesis 5 120

Hypothesis 6 121

Hypothesis 7 121

Summary 122

VIII. FINDINGS RELATED TO THE ANALYSIS OF VARIANCE OF ACHIEVEMENT SCORES 124

The Analysis of Variance 125

Sources of variation in achievement 127

Discussion 133

Hypotheses 139

Hypothesis 1 139

CHAPTER PAGE

Hypothesis 2 139

Hypothesis 3 140

Summary 140

IX. FINDINGS RELATED TO THE ANALYSIS OF VARIANCE OF ACHIEVEMENT FACTOR SCORES 142

Factor Scores 143

The derivation of factor scores 144

The Analysis of Variance of Factor Scores 150

Achievement factor scores 150

Socio-economic factor scores 152

Discussion 154

Total achievement scores and factor achievement scores 154

The size of F-ratios 156

Intraclass correlation 158

Hypotheses 159

Hypothesis 8 159

Hypothesis 9 159

Hypothesis 10 159

Summary 160

X. SUMMARY AND CONCLUSIONS 162

Summary 163

Systems analysis 163

The sample 165

The tests 166

Image analysis 167

CHAPTER PAGE

The analysis of variance of raw achievement scores . . 168

The derivation of factor scores 170

The analysis of variance of achievement factor scores. 170

Conclusions Related to Image Analysis and Principal Axes Factoring 172

Conclusions Related to Scalogram Analysis 174

Conclusions Related to the Analysis of Variance of Achievement Factor Scores 175

Implications of the Study 177

BIBLIOGRAPHY 181

APPENDIX A THE ANALYSIS OF VARIANCE 188

APPENDIX B QUESTIONNAIRES AND RESPONSES 192

APPENDIX C DISTRIBUTION OF ACHIEVEMENT TEST SCORES 209

APPENDIX D DISTRIBUTIONS OF SOCIO-ECONOMIC TEST SCORES 210

APPENDIX E DISTRIBUTIONS OF EDUCATIONAL ABILITY SCORES 212

LIST OF TABLES

TABLE		PAGE
L.	Correlation Matrix of Thirteen Variables	100
II.	The Inverse of the Correlation Matrix of Thirteen Variables	101
III.	The Matrix of Anti-Image Covariance	104
IV.	The Matrix of Image Covariance	106
V.	Eigenvalues and Eigenvectors from the Image Covariance Matrix.	108
VI.	Factor Matrix of Thirteen Principal Axes Factors	110
VII.	Seven Rotated Factors Having Loadings Greater Than .200	112
VIII.	The Analysis of Variance	126
IX.	The Analysis of Variance of Achievement in English Literature.	128
X.	The Analysis of Variance of Achievement in English Language	129
XI.	The Analysis of Variance of Achievement in Social Studies	130
XII.	The Analysis of Variance of Achievement in Mathematics	131
XIII.	The Analysis of Variance of Achievement in Science	132
XIV.	Proportions of Between and Within Variance Associated With Sources in Five Academic Subjects	138
XV.	Inverse Elements of Four Factors of the Matrix A	145
XVI.	The Matrix W	147
XVII.	The Matrix of Weights for Predicting Scores on Four Factors.	148
XVIII.	Means and Standard Deviations of Scores on Four Factors	151
XIX.	Analysis of Variance of Achievement Factor Scores	151
XX.	Analysis of Variance of Socio-Economic Factor Scores	153
XXI.	Distribution of Pupil Responses to Items of a Scale of Attitude to Schooling	202

TABLE PAGE

XXII. Distribution of Scores on the Eleven Item Quasi-Scale of Educational Motivation 205

XXIII. Distribution of Study Hours as Reported by Pupils 206

XXIV. Distributions of Pupil Achievement in Five Grade IX Academic Subjects 209

XXV. Distributions of Scores on the Occupation Scale and the Socio-Economic Status Scale 210

XXVI. Distribution of Scores on the Home Index Questionnaire . . 211

XXVII. Distributions of Scores on Three Tests of Educational Ability 212

LIST OF FIGURES

FIGURE		PAGE
1	A Model of Communication and Control in Systems	29
2	Scalogram of Six Dichotomous Items	91

CHAPTER I

THE PROBLEM

One of the most serious impediments to effective research in education has been the lack of a general theory which would produce hypotheses, suggest experimental designs, and permit the sequential accumulation of results necessary for the generation of new theory. Summaries of recent research in Canadian education are characterized by the wide variety of topics investigated and by their empirical orientation. (Canadian Education Association; Brehaut). Thus in addition to the general controversy which surrounds the conduct of public education (Mayer), there is also a pointed attack upon educational research as such. (Nash). This situation is not limited to the Canadian scene nor to research in education alone.

> In recent years research in the psychosocial sciences has come in for considerable criticism. Adverse comment has been directed against the philosophical preconceptions in terms of which empirical research in these fields has been carried on, against particular imperfections in methodology, and on the grounds of the comparative triviality of both the matters which have received attention and of the results which have accrued. (Bantock, p. 264).

Merton, in examining this same problem in sociology, contends that the most fruitful theory will be of the "middle range" which eschews both the empirically trivial "fact" and the ponderous but untestable "general law." (Merton, p. 9). A pervasive pressure upon all science is one which compels it to keep its activities relevant to the practical world and to produce results of immediate applicability. Even Newton and the Royal Society at the pinnacle of the Age of Reason had to modify their

"disinterested desire to know" to the requirements of shipbuilding, navigation, and the military. (Merton, p. 610). Similar pressures have been exerted upon education as a field of study and often workers attempting to build a social science in education have yielded to the temptation to produce the quick result by generalizing beyond evidence (Cooperative Study; Swanson) or by restricting the problem to particular, non-typical cases. Theory framed in the "middle range" of generality offers little of immediate relevance and is therefore seldom found in education.

Another limitation of educational research is that it is conducted in isolation from the situation in which its findings must be applied. In attempting to replicate the conditions of research in the natural sciences, experiments have been designed to measure the effect of the "single operative variable" by holding all other variables constant. For example, much effort has gone into finding out what makes a "good" teacher and into what methods of teaching reading are most effective. Such studies have served to point out the difficulties of educational research. (Barr; Jenkinson; Anderson). In education, as in all social sciences, variables may rarely be regarded as single and independent. The modification of one variable influences the action of a host of others. For example, it may be hypothesized that there is a relationship between the salary of a teacher and the effectiveness of his teaching. However, higher salaries may make it possible to hire persons with expensive talents--athletic, technical, or artistic. Such persons hired for specific abilities may or may not be capable of carrying out other duties essential to teaching effectiveness. (Kershaw and McKean). Thus it is rarely possible in educational research to believe that we are holding pressure in a gas constant while varying

the temperature to observe the effect on volume.

Moreover, the findings of educational research are rarely related to the cost of applying the findings of such research or to the question of choosing between alternative courses of action. Thus it is not always possible to know that research findings in education are true beyond the particular situation in which they were observed or to know whether their significance justifies the expenditure required to implement them.

These difficulties have been summarized by Kershaw and McKean:

> A great deal of research has of course been done on preferred ways of teaching, organizing schools, using audio-visual materials, and so on. But typically this is in the form of descriptive surveys, in which experience and judgment are the decisive elements, it is "research in small," an examination of a few factors. . . .This kind of research cannot take account of indirect effects (the impact of a changed variable on other variables and thence on output), and there are rarely any costs attached to the alternative practices studied. That the use of phonics makes children learn to read more rapidly is helpful information. But it would be still more helpful to find out whether using phonics to a greater extent, decreasing class size, hiring special teachers, using teacher aides, or buying new electronic equipment would be the best innovation in the sense we discuss here--that is, would achieve the most from a specified expenditure of resources. (p. 4).

Systems in Education

The theory of systems analysis is an attempt to generalize beyond known empirical relationships but without attempting to formulate a "theory of everything." It is thus characteristically one of Merton's middle range theories; it examines whole entities rather than their parts individually; and it regards such wholes as something greater than the sum of their parts. Systems theory may thus answer some of the criticism leveled at educational research. It regards the situation in its entirety so that interdependent variables are accounted for; it permits successive and cumulative research;

it provides a framework for the formulation of hypotheses, and it provides ultimately a rational basis for decision making in educational administration.

A system has three important characteristics. There must be a set of objects having significant relationships between them which operate to produce some outcome. The system has associated with it certain costs which can be expressed in dollars, energy, skills, or labour. There must be alternative ways of combining the elements or objects in the system. (Kershaw and McKean, p. 2).

The applicability of such theory in education is indicated by the fact that we frequently (though imprecisely) refer to "systems of education" such as that of a province or a city. We mean, when we use such a term, the teachers, schools, pupils which, at some cost, interact to produce the outcome learning or education. Clearly the definition of a system is a subjective matter. A classroom may be regarded as a system, or a school, district, or province itself may be the system depending upon the frame of reference which is appropriate. There is virtually no limit to the size of system which may be defined. However, if the theory is to be kept manageable and within the range of "middle theory," as specified by Merton, there must be some knowledge about the interrelationships within the system and some clear indication of their relevance to an output. By specifying the outcome which is of interest the system itself may be defined. If we are interested in knowledge acquired in schools of a selected Canadian province, a system is immediately defined. It may be further specified by stating what kinds of knowledge are to be considered as output and by specifying which objects and which

interrelationships are important and significant. Hall and Fagen (p. 18) state that a room and all the objects in it may be regarded as a system. If that room contains the components of a hi-fi system, these components likely form a more interesting and significant system. When these components are connected and operating a complete system has been defined as objects have been interrelated, they draw energy and information from outside themselves (electricity and a record), and have an output in the form of the sound produced. This example, though not perfect, is a good one since it emphasizes the importance of the objects of a system and their interrelations to the output of the system. One mal-functioning component can alter the output of the system drastically.

Miller offers this definition of systems:

> Systems are bounded regions in space-time, involving energy interchange among their parts, which are associated in functional relationships, and with their environments. General systems theory is a series of related definitions, assumptions, and postulates about all levels of systems from atomic particles through atoms, molecules, crystals, viruses, cells, organs, individuals, small groups, societies, planets, solar systems, and galaxies. . . .
>
> All behavior can be conceived of as energy exchange within an open system or from one such system to another. Any exchange of energy across boundaries results in some alteration or distortion of the energy form. Those specific functions of systems which we can stipulate and whose magnitude we can measure in a relative scale, we will call "variables" if they are within the system. . . . Systems are usually in balance with their environments, which have outputs into systems and inputs from them. (p. 514).

This definition adds the notions of boundaries and environment to the concept of system. It also emphasizes that a system can be an energy exchange mechanism whereby energy is taken from the environment in one form and returned to it in another.

One further notion is required to describe systems in education; this is the notion of sub-system. To return to the example of the hi-fi

system, each of the components of the system may itself be regarded as a system. However, because the output likely to be of greatest interest is sound, the entire assembled mechanism may be regarded as the system and its components as sub-systems.

Though this idea of systems and sub-systems may seem a simple one, it can be of considerable importance for analytical purposes. Once the system and its output have been defined, it is possible to modify relationships between systems, sub-systems, and objects within them. In this way it is possible to evaluate the contribution of sub-systems and objects within them to the output of the system as a whole.

Education is typically organized in a hierarchical fashion which suggests a general system, an environment which provides resources, sub-systems, and elements within sub-systems, and outputs from the system back to the environment. Within the boundaries of the system, it is possible to modify certain relationships and to observe the effects not only upon output but upon other variables and relationships within the system. Thus it is possible to ask in systems analysis, what the effect of increased teacher salaries has, not only upon teaching, but upon hiring policies, upon teacher attitudes, and upon the effectiveness of teaching as well.

> This is what makes the system interesting from an analytical point of view; it is possible to vary the inputs and see what the effect is on output. We can educate children, for example, by having one teacher for each child, or we can do it by using one teacher for each fifty children. Presumably these two different combinations will yield different outputs in terms of the amount of education obtained. . . .In education not only can there be different pupil-teacher ratios, but different salary schedules and levels, different teacher training, building configurations, uses of television, degrees of emphasis on athletics, and so on to infinity. (Kershaw and McKean, p. 3).

Difficulties in Applying Systems Analysis to Education

The application of systems theory in education raises many problems, of course. It is not possible to use schools, teachers, and pupils as elaborate laboratories stocked with useful equipment and guinea pigs. However, it is possible, through multi-variate statistical analysis, to discover relationships between variables within systems and outputs. Such analysis is dependent upon there being differences already existing within school systems. Once relationships are established between these differences and output differences, it is possible to design a model school system incorporating those features which maximize output with a minimum of input. In this way, systems analysis approaches economic analysis. Systems analysis enables maximum efficiency to be gained from limited resources. Or stating the situation conversely, systems analysis can indicate the minimum resources required to achieve a designated amount of output. (Mauch). Some analyses of this type have already been carried out. (Kershaw and McKean; Thomas; Project Talent). Of course, systems analysis has nothing to say about what output is most desirable or about how conflict may be resolved when there are insufficient resources to achieve a desired output.

To an increasing extent, however, educational research is being cast within wider frameworks which use models or systems as reference points. As Bantock states it, effective research in education requires the construction of an "homunculus"--a little man, puppet, or model, which incorporates within it the way we think reality is ordered. The little man is not as complicated nor precise as reality but is easier to manipulate. He may lack the spark of life but he is an integrated unity;

and comparison with the "whole man alive" may suggest means of improving his design and operation. (Bantock, p. 271).

The tendency towards dealing with larger units of reality as integrated units is coupled with the positivistic attitude which insists that reality is that which can be objectified, quantified, and measured. An increasing pressure upon educational research, therefore, demands that the consequences of education be objectified and expressed as a quantitative measurement. Evaluation of the effectiveness of a model or system depends upon such an objective criterion. Otherwise it is impossible to specify how closely our "homunculus" approaches reality; there must be a measure of reality which can be incorporated into the action of the model.

Examples of these trends towards the definition of systematic models and towards the objectification of educational aims and consequences may be seen in studies and reports by such authors as Anderson, Ackerman, Barr, Maccia, Mitzel, Mitzel and Gross, Project Talent, and Thomas. The idealist philosopher may object that important aims in education are not subject to quantification or measurement; or he may claim that even if aims can be objectified, the consequences of educational systems may not be measured until some distant future time when the true effects of present systems become apparent. Educational researchers will not usually wait that long and insist upon available, contemporary measurement. In this way the demands of science affect social values and modify our world.

Another difficulty involved in applying systems analysis to education stems from the diversity and multiplicity of the aims of education. (Enns; Bloom). The complexity of aims in education means that educational

outputs will also be multi-dimensional. Downey lists four broad dimensions within educational objectives: the intellectual, social, personal, and productive dimensions. In a study comparing Canadian and American attitudes to these dimensions, it was found that the Canadian sample assigned high priority to items in the intellectual dimension such as "knowledge, scholarly attitudes, and creative skills." (Downey, p. 44). This attitude is also borne out in the recommendations of recent Canadian Royal Commissions on Education. (British Columbia).

Systems analysis in education can be greatly simplified if output can be defined as a single, objective outcome. The emphasis upon the intellectual dimension in the aims of Canadian education suggests knowledge acquired by pupils as a meaningful and objective measure of output. Systems analyses in education carried out in the United States have also used knowledge of pupils as measured by standardized tests for an output criterion. (Project Talent; Kershaw and McKean).

Designating knowledge as output in a systems analysis of education does not completely solve the problem of what the output of an educational system is. It may also be asked, "Which of the many measures of knowledge is appropriate?" Standardized tests have been questioned on the grounds that they are inadequate measures of knowledge. (Nash). Moreover, knowledge is itself a multi-dimensional output. Even if standardized tests are accepted as a reasonable measure of educational output, it must be recognized that standardized test scores contain measures of intelligence, social status, pupil attitude, and error as well as knowledge. (Elley; Davis; Lord; Tate and Voss). Despite these difficulties, standardized tests are likely the best single measure of educational

output. Their objectivity, relevance to school activities and to later life, and their ready availability all recommend them as measures of educational output. Furthermore, such standardized scores may be modified to allow for effects such as intelligence and social status in order to give a truer picture of knowledge gained as a result of system effects. (Kershaw and McKean, p. 13).

An Example of a System in Education

The organization of education in Alberta offers an example of a system in education which is open to experimental verification in terms of output. This system is organized in terms of several levels or sub-systems. The largest of these sub-system units is the school district or division. One of these units may include a whole city, or, in the case of rural areas, it may include several towns and the area surrounding it. Within each of these districts, various schools form another level or sub-system, and within each school, of course, the classrooms form a third level or sub-system. This system, with its buildings, pupils, teachers, principals, etc., may be thought of as a unit for various reasons. One of the most cogent is that, activities within the system are bent towards shared goals. One of these goals is clearly and rigidly imposed by the requirement that all Grade IX and Grade XII pupils must write examinations administered by the central authority in Alberta education, the Alberta Department of Education. For these two grades, at least, an output has been clearly specified which can be used to define and evaluate the system as a whole. This output is the achievement of pupils on the Departmental tests.

Since achievement is likely to vary from sub-system to sub-system--

that is, between classes, schools, and districts--it is possible to assign the effects on achievement due to each sub-system.

The Purposes of the Study

The major purpose of this study is to apply a systems analysis to selected districts, schools, and classrooms. These districts, schools, and classrooms will be considered to form a system if they can be shown to be related to the achievement of pupils within them. This purpose suggests an analysis of variance design in which variance is assigned to different levels or sub-systems. The question which may then be asked is, "How much of total output variation in achievement is associated with each sub-system?" An analysis of variance of pupil achievement over districts, schools, and classes will permit the breaking down of total output variation into sources from districts, schools, classes, and a residual or pupil source.

A further purpose of the study is to resolve pupil achievement scores into components representing possible sources of achievement variance. Such sources would be the intelligence of the pupil, his socio-economic background, and his interest or motivation, and a factor due to the impact of the system upon learning. The separation of these sources of achievement may be achieved through factor analysis.

The final purpose of the study is to derive scores representing the factored achievement scores. These then may be subjected to the same analysis of variance as the raw acheivement scores to give a clearer picture of the effects of the system.

Definitions

For the purposes of this study, the following definitions are required.

System. A system in education is a group of districts, schools, and classes--together with the resources and personnel which form them--which are interrelated by the fact that they are engaged in producing a similar output.

Sub-systems. Sub-systems are formed within the system as a whole by patterns of activity within it. Thus districts, schools, and classes are sub-systems. Each sub-system has a pattern of activities, resources and personnel characteristic of it.

Elements. The system and sub-systems are composed of elements such as buildings, books, pupils, teachers.

Relationships. Between the elements of the system there are certain significant relationships which impose a meaningful order upon the elements of the system. The administrative characteristics of a district, school, or class are part of these relationships as are teaching methods and the traditional activities of schooling.

Design of the Study

The design of the study involves a hierarchical arrangement of districts, schools, and classes. Twenty-two Alberta school districts participated in the study. In each district, two schools were selected and in each school two classes. Thus an analysis of variance involving twenty-two districts, forty-four schools, and eighty-eight classes was

possible. This selection involved 2,069 pupils for whom complete data were available.

Because of the size of the sample, all data were recorded on I.B.M. cards and most statistical procedures were carried out using the I.B.M. 1620 computer at the University of Alberta.

Analysis of Variance

The analysis of variance design breaks up the total achievement variation as follows:

Total variation = total within variation + total between variation, or in terms of the sub-systems of the educational system in the study:

Total variation = residual variation (pupils) + class variation + school variation + district variation. In this way, F-tests involving a comparison of residual (pupil) variance with class, school, and district variance is possible.

Raw score achievement data for 2,069 Grade IX students were subjected to this analysis. The achievement data were obtained from the examination results of the sample of Grade IX students on the Alberta Departmental Examinations (1962).

Factor Analysis

The achievement data of the student sample were factored using Guttman's image analysis. From this factoring, individual factor scores were obtained and subjected to the analysis of variance as above.

Through factor analysis, the score of an individual on a test may be resolved as follows:

$$Z_{ji} = a_{j1}F_{1i} + a_{j2}F_{2i} + \cdots\cdots + a_{jm}F_{mi}$$

where:

Z_{ji} = the standard score of individual i on test j,

a_{jl} = the loading of test j on factor l,

F_{li} = the factor score of individual i on factor l,

and where common factors run from 1 to m. (Harman; Thurstone; Guttman). The F_{mi} are the factor scores subjected to the analysis of variance, where i runs from 1 to N, the number of students in the sample.

The Intercorrelation Matrix

The intercorrelations of thirteen tests were factored. Five of these were Alberta Departmental Examinations of subject matter knowledge: Language, Literature, Social Studies, Mathematics, and Science. Three were tests of socio-economic status. Two were tests of educational capability: SCAT (Verbal) and SCAT (Quantitative). Two were tests of educational motivation: a scale named Motivation and a self-report by pupils on their study hours. The final one was a test of reading performance.

Hypotheses

The following are the research hypotheses of the study.

H_1: There are significant differences between the mean achievement of classes and the mean achievement of schools.

H_2: There are significant differences between the mean achievement of schools and the mean achievement of districts.

H_3: There are significant differences between the mean achievement of districts and the mean achievement of all pupils.

H_4: The intercorrelation matrix can be factored to yield a factor

on which the Alberta Departmental Examinations of subject knowledge have the highest loadings.

H_5: The Intercorrelation matrix can be factored to yield a factor on which tests of socio-economic status have the highest loadings.

H_6: The intercorrelation matrix can be factored to yield a factor on which tests of educational motivation have the highest loadings.

H_7: The intercorrelation matrix can be factored to yield a factor on which tests of verbal and quantitative ability have the highest loadings.

H_8: There are significant differences between the mean factor score achievement of classes and the mean factor score achievement of schools.

H_9: There are significant differences between the mean factor score achievement of schools and the mean factor score achievement of districts.

H_{10}: There are significant differences between the mean factor score achievement of districts and the mean factor score achievement of all pupils.

Null hypotheses related to analysis of variance. The null hypotheses of the analyses of variance may be stated as:

(1) $\mu_{c_1} = \mu_{c_2} = \mu_{s_1}$

(2) $\mu_{s_1} = \mu_{s_2} = \mu_{D_1}$

(3) $\mu_{D_1} = \mu_{D_2} = \mu_{D} = \mu$

Or in general:

$$\mu_c = \mu_s = \mu_D = \mu$$

where c refers to classes, s to schools, and d to districts.

CHAPTER II

THE NATURE OF SYSTEMS ANALYSIS

Systems analysis has at least two characteristics which make it relevant and generally applicable in the social sciences. In the first place, it deals with complex phenomena as wholes; and secondly it attempts to relate knowledge from various fields into new syntheses. Administration, as a field of study, is also characterized by the complexity of data to be investigated and by the application of knowledge from various sources to the problems encountered. (Griffiths; Lazarsfeld; Halpin, 1958). Indeed, administration may be seen usefully as the study of organization. The nature of organization is one of the complex wholes to which systems analysis may be applied. Many authors, in describing organization, use system concepts such as performance, output, feedback, and control. (Greer; Dubin; Presthus). The nature of systems analysis is thus of considerable importance in social sciences generally and in administration in particular.

General Systems

The concept of general systems and the name "systems analysis" stems from the work of Ludwig von Bertalanffy. As a biologist, he was impressed by the diversity and isolation of knowledge not only in his own field but in all sciences dealing with natural phenomena. He found that science had developed in such a way that models and theories were evolved to deal with areas which were increasingly specialized. In an attempt to

generalize theory, he formulated the concept of general systems.

> There exist models, principles, and laws that apply to generalized systems or their subclasses, irrespective of their particular kind, the nature of their component elements, and the relations of "forces" between them. It seems legitimate to ask for a theory, not of systems of a more or less special kind, but of universal principles applying to systems in general.
>
> In this way we come to postulate a new discipline, called General Systems Theory. Its subject matter is the formulation and derivation of those principles which are valid for "systems" in general. (von Bertalanffy, p. 1).

Thus systems analysis looks for the commonness between orders of data or behavior. If systems of various kinds can be described in similar ways, there may arise a single theory or a general system which accounts for their behavior. General system theorists believe "that the unification of theory in the physical and non-physical world is desirable and ultimately attainable, at least to some degree." (Hearne, p. 38).

The last man with the right or the audacity to claim to have mastered all knowledge may have been Sir Francis Bacon. His synthesis of knowledge, Novum Organum, has not proved to be a useful theory, but systems theory may be regarded as working in the same direction. The obvious pitfall of such an attempt is that in explaining everything, nothing is explained. Systems theorists are aware of this difficulty and attempt to deal with systems which can be operationally defined and which suggest testable hypotheses. (von Bertalanffy; Boulding). Thus whatever their ultimate aim, systems theorists have not approached a general theory of knowledge.

Consequences of Systems Theory

Von Bertalanffy gives two important consequences of systems theory. The first of these is that it permits the "appearance of structural

similarities or isomorphies in different fields." (von Bertalanffy, p. 2). That is, the principles or theories which govern phenomena in one field may be found to be applicable in quite different areas. Thus whenever there are "complexes of elements in interaction," the laws which are found to hold for one complex may hold for others though they may have nothing in common but that they are organizations of behavior. Various examples of such isomorphies are given in the General Systems Yearbook, though most of them are in the general area of the biological sciences. (von Bertalanffy and Rapoport).

As Boulding emphasizes, the social sciences are in need of the application of such isomorphies. Administration has too often been regarded as a technology stemming from legal and economic structure. (Baird). The current trend in administration, however, is to apply the techniques of general social science in an interdisciplinary approach to behavior within organizations. (Downey and Enns). However, the more science is specialized, the more difficult becomes communication between the specialties, and the less likely it is that relevant knowledge will be brought to bear on a problem.

> It is traditional for studies of labor relations, money and banking, and foreign investment to come out of departments of economics. Many of the needed theoretical models and frameworks in these fields, however, do not come out of "economic theory" as this is usually taught, but from sociology, social psychology and cultural anthropology. (Boulding, 1956, p. 12).

The interdisciplinary movement in administration has the potential for drawing useful but disparate knowledge to bear on organizational problems. But systems analysis has the potential of providing a framework for the interdisciplinary movement by providing a general theory and by indicating

areas for useful application of the theory through isomorphies of various kinds. If administration is to be a field of content and not merely the study of rules, it will require a form and structure which can be provided by general systems theory.

Again, Boulding indicates how such theory might be used:

> At a low level of ambition but with a high degree of confidence it aims to point out similarities in the theoretical constructions of different disciplines, where these exist, and to develop theoretical models having applicability to at least two different fields of study. At a higher level of ambition but with perhaps a lower degree of confidence it hopes to develop something like a "spectrum" of theories--a system of systems which may perform the function of a "gestalt" in theoretical construction. Such "gestalts" in special fields have been of great value in directing research towards the gaps which they reveal. Thus the periodic table of elements in chemistry directed research for many decades towards the discovery of unknown elements to fill gaps in the table until the table was completely filled. (Boulding, 1956, p. 11).

A second consequence of systems theory is possibly even more important than the first. It stems from the fact that systems theory can deal with a class of phenomena beyond the reach of classical physical science. Conventional physical science deals only with what von Bertalanffy calls "closed systems." That is, it deals with systems which are isolated from their environments and in which, therefore, there is no interchange of energy between the system and its environment. A closed system may be regarded as a <u>disorganized</u> complexity in which behavior may be explained by "statistics, the laws of chance, and, in the last resort, the second law of thermodynamics." (von Bertalanffy, p. 2). A closed system is illustrated by a sealed vessel in which a chemical reaction is taking place. The final result of the reaction can be predicted to be a state of equilibrium which maximizes de-organization, homogeneity, and the levelling of differences in energy and activity. (Hearne, p. 41).

The result of a reaction in a closed system is the production of entropy, as stated by the second law of thermodynamics.

> In particular, the second principle of thermodynamics states that, in a closed system, a certain quantity, called entropy, must increase to a maximum, and eventually the process comes to a stop at a state of equilibrium. The second principle can be formulated in different ways, one being that entropy is a measure of probability, and so a closed system tends to a probable distribution. The most probable distribution, however, of a mixture, say, of red and blue glass beads, or of molecules having different velocities, is a state of complete disorder; having separated all red beads on one hand, and all blue ones on the other, or having in a closed space, all fast molecules, that is a high temperature on the right side, and all slow ones, a low temperature, at the left, is a highly improbable state of affairs. So the tendency towards maximum entropy or the most probable distribution is the tendency to maximum disorder. (von Bertalanffy, p. 3).

The strength of systems theory is that it can deal with phenomena other than those in closed systems. That is, it can deal with phenomena which are characterized not by disorganization but by organization and in which a dynamic ordering not a state of entropy is attained. Equilibrium in organization implies the ordering of energy and activity as in a living organism rather than the dissipation of energy in random fashion as at the end of a chemical reaction in a closed vessel. The importance of systems theory for the social sciences is that these sciences deal not with behavior which is appropriate to the second law of thermodynamics but with behavior which exhibits wholeness, complexity, purposiveness, and regeneration.

> Today our main problem is that of organized complexity. Concepts like those of organization, wholeness, directiveness, teleology, control, self-regulation, differentiation and the like are alien to conventional physics. However, they pop up everywhere in the biological, behavioral, and social sciences, and are, in fact, indispensable for dealing with living organisms or social groups. Thus, a basic problem posed to modern science is a general theory of organization. General System Theory is in principle capable of giving exact definitions for such concepts and, in suitable cases, of putting them to quantitative analysis. (von Bertalanffy, p. 2).

Characteristics of Systems

The most concise definition of a system has been given by Hall and Fagen (p. 18). "A system is a set of objects together with relationships between the objects and between their attributes." To make this definition more meaningful, certain other definitions and characteristics of systems should be explained.

Objects. The objects in a system are its parts, elements, or components. Such objects are usually physical; in an educational system, objects are schools, books, teachers, pupils, and etc.

Attributes. Attributes are properties of objects. Attributes of schools would be their size, age, esprit, facilities, and so forth. Attributes of teachers would include their training, knowledge, experience, salary, and sex. Similarly pupils have the attributes of intelligence, social status, and motivation.

Relationships. The relationships referred to are those which "tie the system together." They are the relationships which define the system and make it significant or important for some purpose. In school systems, relationships hold between economic resources of a district, the administration of schools, the competence of teachers, the attitudes of pupils, and the output of knowledge gained by pupils. There are, of course, a multitude of relationships which might hold within systems. Trivial relationships may be distinguished from significant ones depending upon the problem set for the system analysis or upon the purposes of the system. (Hall and Fagen, p. 18).

Boundary. Systems are bounded in the sense that objects may be either in or out of the system. A university would not usually be thought of as part of a public school system even though it is an institution of learning. That is to say, a university is not part of the relationships which influence the learning which go on in, say, a Grade VI classroom. However, a university could be part of a larger educational system if it has responsibility for training teachers or if it made efforts to influence the standards or practices of public education. Thus boundaries, like relationships in systems, are determined by the nature or purposes of the system.

Boundaries may be operationally defined in terms of the intensity of interaction within groups of objects (Hearn, p. 42) or in terms of energy concentrations. (Miller, p. 514). In both cases the definition may be somewhat arbitrary, though open to later empirical validation.

Environment. The environment of a system is everything beyond its boundaries which affects or is affected by the operation of the system.

> As in any scientific activity, one includes in the universe of system and environment all those objects which he feels are the most important, describes the inter-relationships as thoroughly as possible and pays closest attention to those attributes of most interest, neglecting those attributes which do not play essential roles. (Hall and Fagen, p. 20).

What the boundaries, objects, and environment of a system are depends on the unit of activity to be analyzed. Thus a record may be part of a hi-fi system or it may be regarded as part of its environment. Similarly, universities, private schools, and public schools may be thought of as parts of one general educational system, or they may be regarded as separate systems.

Sub-systems. From the definitions of environment and boundaries, it is evident that systems are hierarchical arrangements of smaller systems, with the definition of the system depending on the unit of analysis. If we are dealing with a general educational system, universities, public and private schools are sub-systems within a larger system. "Every order of system with the exception of the smallest has sub-systems, and all but the largest are part of a supra-system consisting of the system and its environment. . . ." (Hearn, p. 41).

Open and closed systems. Systems may also be described in terms of their openness or closedness. Closed systems have already been described as those to which the second law of thermodynamics is applicable. Such systems are isolated from their environments; laws which hold for one closed system hold for other similar closed systems. Part of the difficulties of educational research, as described by Bantock, stem from an attempt to regard the data of education as part of closed systems when in fact these data are very much altered by the nature of the environment in which they occur. The well-known Hawthorne studies at the Western Electric Company demonstrated clearly that the effects of seemingly closed system variables such as intensity of light, frequency of rest periods, and methods of work were completely engulfed by the environmental effects of worker attitude and worker organization.

General systems analysis deals with open systems rather than closed. It is this characteristic, as von Bertalanffy points out, which makes it so useful in the social and biological sciences. Open systems are those which exchange matter, energy, or information with their

environments. They are characterized by organization rather than disorganization as is the case in a closed system in equilibrium. Open systems do not fit the postulates of conventional physical science; it was this consideration which led to the development of general system theory.

> . . .We find systems which by their very nature and definition are not closed systems. Every living organism is essentially an open system. It maintains itself in a continuous inflow and outflow, building up and breaking down components, never being, so long as it is alive, in a state of chemical and thermodynamic equilibrium but maintained in a so-called steady state which is distant from the latter. . . . (von Bertalanffy, p. 3).

Steady state. Steady state is a condition reached in open systems. It resembles the equilibrium reached in a closed system but is different from it because it represents dynamic organization rather than entropy or disorganization. A candle, immediately after it is lit, or an engine after it is started, performs differently than after it has warmed up. The flame at first is small and blue; the engine is noisy or erratic. Finally these systems attain a typical state of organization which is their steady state. This state persists despite minor environmental changes until fuel or oxygen are consumed. It is as though the steady state were the aim or purpose of the system. White rats, faced with increasing stress through environmental changes (such as rapid temperature changes), maintain their characteristic steady state (activity, body temperature, etc.) until some critical value is reached in the environment, and then die in a "catastrophic collapse" of their steady state. (Hearn, p. 68).

Social organizations, too, have their steady states (Dubin; Greer;

Presthus) and resist modification through changes in environment. Restriction of output in work groups is an example of steady state in social organization. (Homans). This tendency to maintain a steady state which is characteristic of open systems von Bertalanffy calls "negative entropy". It is a measure of order, organization, or improbability. By taking "negative entropy" into account, general systems theory can resolve the paradoxes of open systems. (von Bertalanffy, p. 5).

> Both closed and open systems are capable of attaining stationary states, the nature of which, however, is different, in each case. A closed system must eventually reach a state of equilibrium. . . . An open system may attain a stationary state in which the system appears also to be constant although maintaining its constancy in a continuous change, inflow, and outflow of materials. (Hearn, p. 41).

A school in a steady state would be typified by the maintenance of certain activities, attitudes, and levels of output. An implication to be drawn from the hypotheses of this study is that, besides the steady state characteristic of the system under investigation as a whole, each of its sub-systems, as seen in districts, schools, and classes, will have its own typical steady state. These steady states of the sub-systems, although similar, will vary sufficiently to establish significantly different levels of output.

Equifinality. Closely related to the concept of steady state is that of equifinality. It is one of the paradoxes of open systems. Equifinal conditions are attained when different or varying inputs into a system result in identical output. In contrast the processes occurring in closed systems are irreversible; with given inputs into a system, the results are inevitable and completely predictable.

In any closed system, the final state is unequivocally determined

> by the initial conditions; for example, the motion in a planetary system where the positions of the planets at a time t_t are unequivocally determined by their positions at a time t_o. . . .If either the initial conditions or the process is altered, the final state will also be changed. This is not so in open systems. Here, the same final state may be reached from different initial conditions and in different ways. (von Bertalanffy, p. 4).

As an example of equifinality in biology, von Bertalanffy cites the case of the sea urchin which can develop from a single ovum, half an ovum, or the fusion of two ova. In man, identical twins result from the splitting of a single ovum. (von Bertalanffy, p. 5). Paradoxes such as these, have led some biologists to explain these phenomena in terms of soul-like, vitalistic notions. However, in general systems theory, the concepts of steady-state and equifinality offer explanations of seemingly purposive behavior which avoid the non-rational approach of the vitalists and the "supposed violation of physical laws." That is, given certain critical values or predisposing conditions in the system and its environment, systems are set upon one course or another which may lead to quite different, though functioning, steady states. This explanation is similar to Mead's social-psychological explanation of "mind." The life of man with a brain that has no symbols in it is quite different from that of symbol-manipulating man, as is apparent from the evidence gained from the observation of feral children. The physical properties of the brain of man with symbols in it and the brain of man without symbols may be identical but their characteristic steady states are totally different. The difference is that between civilization and animal existence. Thus the presence of the non-physical symbol is a critical value altering completely the operation of the human system.

Equifinality has an important implication for educational systems.

Given that schools, classes, and districts have characteristic steady states, these sub-systems will function to produce constant outputs. Thus we can conceive, for example, of a classroom functioning to maintain output at a high level despite varying inputs such as pupils of different intelligence, social status, and attitude. Or conversely, there could be schools whose steady state acted to depress output despite its new building, elaborate equipment, and ample economic resources. The equifinal process in systems would predict, too, that educational systems with similar inputs could achieve different outputs.

Feedback. The seeming purposiveness of systems is due to the operation of feedback mechanisms within them. Feedback occurs when portions of the output of a system are re-directed back into the system as input. (Hall and Fagen, p. 23). Or to define it in a way which emphasizes the control and purposive aspects of system functioning, Wiener describes feedback as follows:

> Feedback is the property of being able to adjust future conduct by past performance. Feedback may be as simple as that of the common reflex, or it may be a higher order feedback, in which past experience is used not only to regulate specific movements, but also whole policies of behavior. . . .
>
> For all these forms of behavior, and particularly for the more complicated ones, we must have central decision organs which determine what the machine is to do next on the basis of information fed back into it. . . .
>
> It is easy to make a simple machine which will run toward the light or run away from it, and if such machines also contain lights of their own, a number of them together will show complicated forms of social behavior. . . . (Wiener, p. 33).

The operation of a thermostat is an example of a system with feedback. Feedback is not limited to machines as it occurs in both natural and man-made systems. The operation of feedback mechanism explains how

systems are enabled to attain steady states. The homeostatic balance in living organisms and the operation of a computer are both examples of complex systems dependent upon feedback.

The model of feedback in a system is given in Figure 1. (von Bertalanffy, p. 5). This diagram emphasizes the fact that control without feedback is impossible. In social organizations, control rests largely with administrators. (Halpin, 1957). If control is inadequate, there may be several reasons for this inadequacy--most of them connected with feedback mechanisms. First, there must be a selection or sampling of output in a "quality control" sense. Secondly, the control mechanism (the administrator) must have receptors capable of receiving the feedback. Third, the control mechanism must be able to interpret feedback received and be able to make decisions in accordance with it. Finally, the control mechanism must feed orders to effectors which are capable of implementing the decisions of the control mechanism. Feedback may also be regarded as information and, in this sense, systems are dependent upon a flow of information. Any flaws in receptor, control, or effector mechanisms can destroy the flow of information or render it useless. (Boulding, 1953).

In bureaucratic organizations such as educational systems, feedback may be inadequate for various reasons. Proliferation of functions and increase in size makes the flow of information more difficult and administrators become increasingly isolated from the information which they need to make decisions. In this case, as Boulding says, channels of communication, which the administrator thinks of as windows become more like mirrors. Even if communication is clear, the personality of

the administrator may render him incapable of making effective decisions. Or as is more often the case, the administrator simply does not know what effectors are appropriate to carry out certain decisions. If reading achievement in a school is low, can we be immediately sure of what causes this state or of what will increase it?

All of these problems can be made clearer and courses of remedial action suggested through the application of systems analysis. That is, if aims, means, and output are consciously related in a system, an increase of information about the system is possible and means of optimizing output may be suggested. Since feedback is information and information is the negative entropy which maintains organization, it is feedback

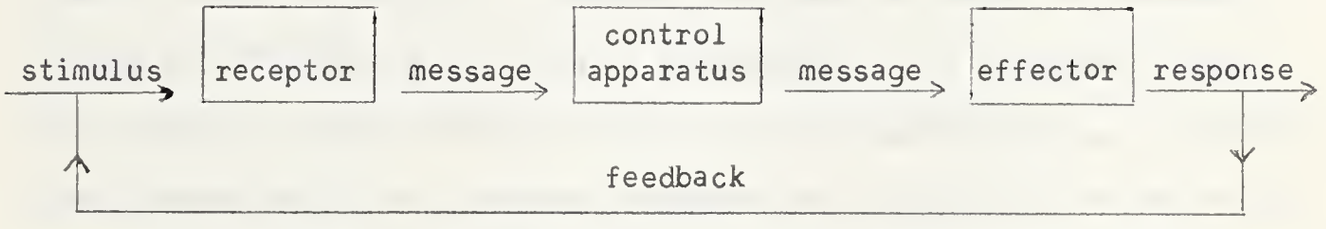

FIGURE 1

A MODEL OF COMMUNICATION AND CONTROL IN SYSTEMS

(or the lack of it) which maintains educational systems in their characteristic steady states.

Wholeness and independence. If systems are so arranged that a change in one part of it affects other parts, it is said to have wholeness. Similarly, at the other end of this continuum, systems are said to be independent if a change in one part has relatively little effect on others. A system is said to be undergoing progressive systematization if it is

moving towards wholeness and to be undergoing progressive segregation if it is moving towards independence. With either progressive systematization or segregation a system may simultaneously be undergoing progressive centralization. A system is said to be centralized when "one element or sub-system plays a dominant role in the operation of the system." (Hall and Fagen, p. 22). The operation of the human body is an example of a relatively independent system with centralized control.

Optimization. Systems frequently operate not only to adapt to their environment but to take the maximum advantage of them. When feeding and water conditions are favourable, animal populations increase accordingly and when demand is high, business organizations increase in number, size, and activity. Systems also tend to strive for maximum output for minimum input. This is optimization. In man-made systems, this tendency creates difficult decisions for administrators. In telephone transmission, the optimum bandwidth for transmission of all tone qualities is not the same as the optimum width for economic (inexpensive) transmission. (Hall and Fagen, p. 24). In educational systems, one method of optimizing output would be to educate only a select group of students. Such a policy however, may not be justified on philosophical and political grounds. Thus systems analysis cannot be regarded as an infallible guide to the optimization of output.

Systems and Organization

Much current literature in administration is concerned with establishing a theory of organization. (Merton; Walton; Bidwell; Dubin; Presthus; Boulding, 1953). In fact, as Boulding suggests, "the idea that

a theory of organization is possible in one of the most important ideas of our time." (Boulding, 1953, p. 67). Bidwell suggests that administrative activity can be divided into three concerns--the sociological, political, and economic. The sociological questions concern the influences of social groupings upon the individual and the effects of such groupings upon the organization itself. This question is examined by Dubin and Presthus and others concerned with "human relations" in administration. Economics enter administration by virtue of the fact that most organizations have explicit purposes and set about to achieve them through the expenditure of resources. Economics thus influence the structure of organizations by imposing ends-means chains and by creating a rational, and hierarchic structure. Political questions enter administration since economic activities require the allocation of scarce resources, choices between alternative courses of action, and authority for making decisions.

One of the few authors to attempt to state a general theory of organization which includes all these aspects is Boulding. It is his stated purpose to apply the theory of general systems as found in biology to human organization and to administrative systems in particular. The similarity between biological and social organization is established by their both being bounded complexities, by their exchange of input and output with their environment, and by cyclical patterns--birth, growth, maturity, and death--in their activities. The aim of a theory of human organization would be to explain not only these cyclical regularities but the "organizational revolution itself"-- the startling growth in modern times of the number, size, and complexity of organizations.

The approach to these questions taken by Boulding is to regard society as an "ecosystem," a macrocosm of variables in which living systems grow, compete, expand, or shrivel and die. One of the simplest examples of an "ecosystem" is a balanced aquarium.

> Society likewise is a great pond, with populations of families, grocery stores, gas stations, automobiles, countries, churhces, lodges, universities, and innumerable other species of social organizations. If we are to understand the laws which govern the growth and survival of any class of organization, we must understand where it fits into the social ecosystem. (Boulding, 1953, p. xxi).

Within ecosystems there are certain critical values which permit development of one kind and suppress those of another. On one side of these values, certain organizations may develop and on the other quite different ones. The climatic differences between the grassy, treeless plains of the Prairies and the lush eastern forests is quite small but great enough to produce totally different biological systems.

> The coexistence of different ecosystems can be explained by quite slight changes in the critical variables. . . .In society many widely differing social and ideological ecosystems coexist in the same region: the lush forest of Roman Catholicism and the open prairie of Boston Unitarianism, for instance, existing side by side in different microclimates or social strata of the same society. (Boulding, 1953, p. xxv).

The importance of organizations is that they do things; the aim of human organizations is to accomplish purposes of people and to enable them to live better lives. Educational organizations are doubly important in that they not only do things for people, they do things to them. Our concern is that they should do the right things and do them well. The ethical implication of a theory of organization is thus made clear in that it enables us to ask why organizations behave as they do or why they fail to perform in expected ways. Boulding makes clear the relevance of

organizational theory to ethical problems in this fervent statement:

> Why is human progress so slow, and so self-defeating? Why, in spite of all the ethical insights and vision of the prophets and saints, does man still dwelll in poverty and insecurity, pride and lust? Why do his civilizations come to grief, and why does his technical progress exhibit such ups and downs? Why do his achievements turn to dust, and his very successes drive him all the faster to destruction? We know indeed that the spirit giveth life. But how is it that the letter killeth? Why do all our attempts to <u>organize</u> great moral insights turn out so badly: the Church of the loving Saviour producing the Grand Inquisitor, Communism becoming enshrined in a vast instrument of exploitation, democracy and science working hand in hand to produce the horror of Hiroshima?
>
> The answer is not in the inadequacy of our ethical ideas. . . . The trouble lies, of course, in the "flesh"; that is, in organization. . . .All organizations possess laws of their own being which they impose on their members, laws which may not be consistent with the purposes of these members. . . .Unless we understand the laws of organization, at least in some degree, we cannot hope to create organizations which will serve the ends of man instead of frustrating them. (Boulding, 1953, pp. 66-67).

We may regard organization as (in Boulding's terms) the apparatus for the "righting of wrongs," or we may more coolly regard it as a means of getting the most for our money. In either case, we must know how the apparatus works, if we are to achieve human purposes. General systems theory provides not only a general framework for the examination of organization but a group of concepts from the cybernetic model that can be operationalized into the functions of organization. Organizations, like thermostats, are mechanisms of control and we can ask through systems analysis to gain better command of that control.

Society is usually quick to put to use knowledge discovered by science; and the problems with which science must concern itself are greatly influenced by society. The gap between organizational aims and achievement is one of the critical problems of contemporary life. "In this regard the scientist and the saint have strictly complementary tasks:

the scientist to show where we are and the saint to show where we ought to be." (Boulding, 1953, p. 73).

There are thus two aspects to the problem of organization: the technical and the ethical. The officers of the White Star Line who struck the staff of the Titanic off their payroll from the moment the ship disappeared below the surface of the sea (Bidwell) had a limited view of their technical function and none of their ethical role. In the final analysis, however, these two aspects may well be combined, since an immoral but technically competent organization would defeat human purposes.

The Iron Laws of Organization

Boulding lists seven "iron laws" of organization in an effort to outline the general behavior of organizations. These "laws" are generalizations about organizations suggested by the systems approach rather than the results of systematic experimentation. However, they do suggest hypotheses for further research. (Boulding, 1953, pp. 77-80).

The Malthusian law of growth. The Malthusian law states that "if there are no checks on the growth of population except starvation and misery, then the population will grow until the people are miserable and starve." This law applied to organizations suggests that there is an "increasingly unfavourable environment" as organizations grow. Cyclical fluctuations in animal populations are an example of this law in nature. Political, business, and religious organizations are affected by it also. A business organization marketing a new and superior product may find its enterprise grows rapidly until larger and (perhaps) more ruthless competitors are affected, at which time they will move to restrict or

negate the advantage of the new enterprise. The growth of the Communist party in the United States or of the sect of Jehovah's Witnesses in Quebec is an example of growing organizations which very quickly felt the restrictions of environmental conditions.

The law of culture patterns and economic activity. Culture patterns and economic activity are closely related as is evidenced by the relationship of Protestantism to Capitalism or by the poverty of India to its timeless, other-worldly culture. Since culture is embodied in familial relationships, permanent social changes must also involve changes in the family.

The law of optimum size. Organizations have optimum sizes which may be large or small depending on their purposes and structure. Changes in size away from the optimum will alter the effectiveness of organizations. Growth in size in all organizations creates increasing difficulties in communication and makes relationships within an organization more impersonal and coercive. For example, it is apparent that large universities are different from small and that a class of one hundred and fifty students is qualitatively different from a class of ten. What the optimum size of classes, schools, or universities is, is still a matter of conjecture.

The law of hierarchy. The law of optimum size has a corollary in the law of hierarchy which states that increasing size in an organization creates more differentiations and sub-divisions within it. With increasing complexity within the structure of an organization, communication becomes more difficult. Communication in administrative

organizations is frequently better organized downwards than upwards. (Dubin). That is, the control mechanism has effective means of transmitting orders and of influencing the behavior of the numerous sub-systems of the organization, but at the same time has difficulty in learning what the effects of such orders are. Therefore the basis on which decisions are to be made in the control mechanism becomes obscured. In the terms used by Hall and Fagen, growth of organizations is accompanied by progressive segregation and progressive centralization.

The failure of communication in organization is a basic cause of inefficiency or inadequacy in organization. Systems or organizations, as has been emphasized, are instruments of correction. Correction cannot be made if there is no information about what needs correcting or about the effects of previous attempts at correction.

The law of oligopoly. This law states that when the number of competing organizations is few, " a situation of acute instability and conflict will be created." The instability and conflict come about because a small number of organizations in an ecosystem increases the effect which a change in any one will have upon the others. That is, the power of each organization over the others is maximized. Current international tensions may be seen to stem from conflict between the two great world powers, the United States and the U.S.S.R.

The law of instability in uncontrolled interaction. Not only does a reduction in the number of competing organizations create instability, but the "uncontrolled interaction" of large numbers of independent organizations may also create instability. Stock market panics, inflation

and deflation, fads, food surpluses and hoarding are examples of this law.

The law of role persistence. Organizations are made up of groups of roles linked in networks of communication. The law of role persistence states that in the interaction between role and incumbent, the role dominates. This is so not only because of the coercion of the particular role held by an individual but because of the pressures of all other interacting roles. There are, of course, pressures from the individual to modify the role. "A square peg in a round hole will do something to round the hole." The questions which arise, then, are when to fit an incumbent to a role and how to modify a role when modification is necessary.

Communication and System

The question of communication within systems has already been indicated as being of critical importance. Moreover, communication increases in difficulty as the organization expands and becomes more complex. The failure of organization, then, is largely due to the failure of communication. What causes such failure? Using the basic cybernetic model (Figure 1), Boulding indicates three causes.

Failure in receptors. Receptors may fail because they do not read or do not receive either stimulus or feedback. In educational systems, the receptors are all the personnel or instruments maintained for the measurement of output and input. These receptors, if working accurately, would indicate the abilities and needs of pupils entering the system and would indicate the changes brought about in pupils as a result of system

activity. An example of a receptor in a physical system which is often given is the thermostat. To perform its function, the thermostat must perceive both real and ideal conditions; it must perceive the actual present temperature and the desired temperature. In education our ideals are often vague (Enns, 1962) and the measurement of them difficult.

Failure in transmission. Even if the receptors read both the real and the ideal accurately, they must be able to communicate their information to the executive or control mechanism. As organizations become more diversified and centralized, communication upwards towards the centers of control becomes slower and more difficult.

> A communications system from the receptors to the executive is essential if an organization is to continue. . . .Otherwise the executive does not receive a realistic picture of his environment, and especially does not know whether or not his actions are taking effect. He creates, as it were, a spurious environment of his own--an environment, if he is a dominant type, of yes-men who reflect the executive's own wishes rather than the realities of the world around him. . . . (Boulding, 1953, p. 70).

Failure of the executive. Having received an accurate picture of the environment, the executive must then transform this information into decisions which are then transferred to effectors. Thus the executive must be able to translate perceptions into appropriate action; and he must have at his command a battery of specialized effectors which can be manipulated to operate accurately for specific purposes. In this way, the executive knowing what is "right", gives instructions to the effectors that will produce "better" information from the receptors. (Boulding, 1953, p. 70).

In educational systems, a detailed and continuous flow of information

about the precise state of the system is not always available. One of the reasons for this situation is that neither aim nor output has been clearly specified. School budgets are object classifications which "show what education buys (teacher time, books, chairs) not what it does (teach maths, science, history)." (Mauch, p. 159). A "performance budget" of the kind which would be consistent with this model of systems would "assign costs to major inputs in combinations that carry out tasks or functions." (March, p. 160). In all systems analysis, the quantification of variables, the measurement of input and output, and a continuous flow of information are essential for approaching rational decision making and for efficient operation of the system.

Applications of Systems Analysis in Education

Actual applications of systems analysis in education are very few in number. Kershaw and McKean have constructed a model for systems analysis which suggests the difficulties and costs of making anything like a complete systems analysis. Their account indicates specific methodology and applications of systems analysis; it also provides sample data and hypothetical results. They have indicated the general application of systems analysis in education as follows:

> In a successful systems analysis, the analyst can vary the inputs, note the effect on both cost and output, and then decide that one system is better than another. He can tell you that if you have a certain sum of money and if you agree with him on what is to be produced, your money is best spent on a combination of inputs he can specify because this combination will maximize output. Or, what is logically equivalent, if you tell him what output you want, he can tell you how to get it at minimum cost. (Kershaw and McKean, p. 3).

Besides the definition of the problem and the indication of a methodology for testing hypotheses about systems, Kershaw and McKean do not offer

concrete findings about education systems.

A study which employs the cybernetic systems model to problems of classroom instruction is the one described by Bushnell and Cogswell. In this study a computer was used to regulate the flow of information to pupils so that individual learning situations could be presented to each pupil in the class. Control of rates of learning was maintained jointly by the computer and a teacher, with the teacher having general supervisory control and the computer greatly increasing the flow of information about rates of learning and providing each pupil with a lesson appropriate to his skill.

The most ambitious and complete application of systems analysis to education is given by the study known as Project Talent. This project involved a stratified sample of 450,000 high school students throughout the United States. In it, numerous measures of educational output were related to characteristics of pupils, schools, and communities. Dailey reports, from this study, a correlation of .876 between a total information score and per pupil expenditures. (Dailey, 1962). From the same study, a correlation .880 is reported between the total information score and per capita income. Considerations such as this led to the classification of schools in the sample. It was found, for instance that the location of the school was one of the most important factors influencing output. The differences in location also represent differences in socio-economic status, economic resources, and attitudes among the locations.

> Other things being equal, location is an important factor influencing the results obtained by a given school. In some locations serving homes in favored environments, it is much easier to obtain high achievement and similar education results than it is in other locations serving homes in unfavorable environments. This shows up

> dramatically in our largest cities. The achievement levels of the lowest and highest academic high schools within any one of our largest cities differ far more than do the average achievement levels of the highest and lowest Regions of the country. the reasons for this are many and complex, but most of them are <u>not</u> under the control of the principal and the staff of a given school. (Dailey, 1962, p. 1).

The findings reported by Dailey also indicate some of the difficulties of systems analysis due to spurious correlations. For example, the relationship of school size to location or region is very high, giving a spuriously high correlation between size of school and achievement of pupils. When region is held constant, Dailey reports that the relationship between total scores in English and size of school is virtually zero. (Dailey, 1962, p. 13).

> Therefore, when schools of different sizes are compared directly in relation to school outcomes, the differences are largely really a reflection of regional differences. Regional differences in turn are largely reflections of rural-urban and socio-economic factors. When these are held constant, regional differences are greatly reduced. (Dailey, 1962, p. 2).

The findings of Project Talent as reported by Dailey raise the questions of precisely what differences in output are due to system variations and of what variables within systems produce these differences. Some light is thrown on these questions by an extension of Project Talent completed and reported by Thomas. This study involved 206 high schools of Project Talent selected to form a stratified random sample of U. S. schools. Multiple regressions were used to relate thirty-two independent variables in schools and communities to eighteen dependent variables based on pupil achievement. The thirty-two independent variables represented financial inputs, organizational arrangements, and socio-economic characteristics of the community; the dependent variables were obtained

from the eighteen tests of an achievement test battery. (Thomas). It was found that the thirty-two variables representing school, home, and community characteristics accounted for seventy-six per cent of variance in the Test of Total Information. The multiple correlation reported for this test is .8699. Twelve variables accounting for the greatest variation in achievement scores were identified using step-wise regression. These variables were then presented in the order determined by the regression:

> That is, the first is that which contributed most to the explanation of the dependent variable's variance; the second contributed most to the remaining unexplained variance, and so on. The direction of the relationship is that which obtains when the preceding variables have been held constant. (Thomas).

The twelve most highly predictive variables and their directions are given below:

1.	Mean starting salary--male teachers	Positive
2.	Quality of housing in the area served by the school (the negative correlation indicates that high achievement scores are related to good housing)	Positive
3.	Percentage of boys in last year's graduating class who went to college	Positive
4.	Age of school building	Positive
5.	Median family income	Positive
6.	Average daily percentage of absenteeism	Negative
7.	Number of school books in school library	Positive
8.	Type of secondary school (the positive relationship here indicates that higher achievement scores are obtained in comprehensive schools)	Positive

9.	Does the school have a guidance program	Positive
10.	Size of community (population)	Positive
11.	Experience of teaching staff	Positive
12.	Grades included in the secondary school (the schools that include from Kindergarten to Grade 12 or Grade 1 to Grade 12 do more poorly than all other schools combined)	Negative

The high predictive value of these twelve variables held for all eighteen achievement variables with the order of them being relatively constant. (The order given here is for the test of Total Information). A significant exception occurred in the prediction of Grade 12 English where "starting salaries fell far behind, and the most powerful predictors were quality of housing and, significantly, the number of books in the high school library." (Thomas).

Summary

This chapter has outlined the framework of general systems theory and has summarized the major concepts involved in it. It has also indicated the relevance of systems theory to the study of organization and has presented the findings of two large studies dealing with education in terms of general systems theory. The purpose of the present study is to measure the effects of broad organizational arrangements in a specific educational system upon output. Raw score pupil achievement is to be resolved into system effects and pupil effects. A further aim is to derive an achievement score free of socio-economic and intelligence components and to subject these scores to the same analysis as the raw

achievement scores.

The four succeeding chapters deal with the theoretical framework for the treatment of data, the collection of data, the test data and the analysis of the data. The next three chapters deal with the findings of the study; and the final chapter presents the conclusions of the study.

CHAPTER III

THEORETICAL FRAMEWORK FOR THE ANALYSIS OF DATA

The data of the study were analyzed using the following statistical techniques: analysis of variance, image analysis (a variation of factor analysis), derivation of factor scores, and scalogram analysis. This chapter outlines the application made of each of these techniques. Also included are discussions of problems in the measurement of socio-economic status and educational ability.

Analysis of Variance

The systems analysis approach of the study suggests an hierarchical arrangement of sub-systems involving school districts, schools, classrooms, and pupils within classes. In each of the twenty-two school districts participating in the study, two schools were selected, and in each school, two Grade IX classrooms were used. For each pupil, data involving achievement on Alberta Departmental Examinations, socio-economic status, academic ability, motivation, and hours of study were collected. One of the major purposes of the study is to assign the proportion of academic achievement associated with major sub-systems such as district, school, and class. This hierarchical arrangement of pupils within districts, schools, and classes suggests, in turn, an analysis of variance design. The design adopted may be considered a modified three-way analysis of variance with missing blocks or as a partition of total variation into

within variation and between variation in which the total between variation is further divided to represent districts, schools, and classes as sources of between variance. The model for the analysis of variance is as follows[1]:

Let i = 1,...., 22 districts,

j = 1, 2 schools in a district,

k = 1, 2 classes in a school,

L = 1, ..., N students in a class.

Then S_t (the total sum of squares) may be partitioned into S_r (the residual sum of squares), S_c (the between classes sum of squares), S_s (the between schools sum of squares), and S_d (the between districts sum of squares) as follows:

$$S_t = S_r + S_c + S_s + S_d$$

where:

$$S_t = \sum_i \sum_j \sum_k \sum_L (X_{ijkL} - \bar{X})^2$$

$$S_r = \sum_i \sum_j \sum_k \sum_L (X_{ijkL} - \bar{X}_{ijk})^2$$

$$S_c = \sum_i \sum_j \sum_k N_{ijk}(\bar{X}_{ijk} - \bar{X}_{ij})^2$$

$$S_s = \sum_i \sum_j N_{ij}(\bar{X}_{ij} - \bar{X}_i)^2$$

$$S_d = \sum_i N_i(\bar{X}_i - \bar{X})^2$$

Thus total variation arises from differences between the achievement of each pupil and the grand mean of achievement in the sample.

---

[1]For the derivation of this model, the author is indebted to Dr. J. R. McGregor of the Department of Mathematics, University of Alberta.

Residual variation arises from differences between pupil achievement in a class and the mean achievement of a class. Class variation arises from differences between class means and school means. School variation arises from differences between school means and district means. District variation arises from differences between district means and the grand mean. If the means of classes, schools, and districts are similar to each other and to the grand achievement mean, the only contribution to between variance is the difference between pupil achievement and class means. In this case, all F-tests will be non-significant.

A complete presentation of this model with additional definition of notation and computation formulas is given in Appendix A.

From the partition of sums of squares, estimates of variances for between classes, between schools, between districts, and the residual are possible. F-tests on sources of variance may be constructed as follows:

$F_{classes}$ = Mean square between classes/Mean square residual,

$F_{schools}$ = Mean square between schools/Mean square residual,

$F_{districts}$ = Mean square between districts/Mean square residual.

Image Analysis

Image analysis is a modification of factor analysis developed by Guttman to meet the problems of the indeterminacy of factors found in standard factor analysis. (Guttman, 1940; 1953; 1954; 1956; 1958). The modification is recommended not only because the additional calculations are easily made but also because the consequences of it are of theoretical and practical importance. (Kaiser).

The aim of factor analysis is to "explain" the intercorrelation of variables in terms of components or factors within the variables. Spearman and his colleagues in England at the turn of the century observed that matrices of intercorrelations of achievement data tended to have columns proportional to each other and to yield zero or positive intercorrelations.

> Then the hypothesis is that there exists a single component c--or one general factor--such that if it were eliminated, then the inter-test partial correlation would be zero:
>
> $$r_{jk.c} = 0$$
>
> . . .The correlation of r_{jc} of a test $_j$ with the component c is called the common-factor loading of the test. Using this language, Spearman's basic equation
>
> $$r_{jk} = r_{jc}\, r_{kc} \quad (j \neq k)$$
>
> says that the correlation between any two tests t_j and t_k is the product of their common factor loadings. (Guttman, 1954, pp. 261-263).

Since no empirical data could be found to meet these hypotheses exactly, other psychologists (notably Thurstone in the United States) hypothesized not one component but many in test intercorrelations. Thus multiple-factor analysis was born. The Thurstonian hypotheses are similar to Spearman's except that test intercorrelations are explained by multiple factors rather than a single one.

Unfortunately, the Thurstone technique introduces problems not encountered in the Spearman model. The Thurstone hypotheses require that either the number of factors or the communalities--the sum of the squares of factor loadings for each test--be known in advance of the analysis. However, there is no <u>a priori</u> knowledge of the number of common factors or the communalities and one cannot be calculated without the other. (Harman, pp. 69-96). In the face of this difficulty, multiple factor analysis has proceeded using approximation methods and various tests for

the significance of factors.

> The usual computing routines operate like a sort of "grinding machine" for grinding out common factors. If one common-factor doesn't suffice to reproduce the r_{jk}, grind out another. If these two together don't suffice, grind out a third, etc., until the approximation to the r_{jk} is regarded as adequate. (Guttman, 1954, p. 266).

Both factor analysis and image analysis arrive at the following important consequences:

$$Z_j = a_{j1}F_1 + a_{j2}F_2 + \cdots + a_{jm}F_m + a_jU_j, \quad \text{(Harman, p. 13)},$$

that is, that a test Z_j may be broken into components representing loadings of the test (the a_{jm}) on common factors, F_m, and the loading of the test on a unique factor, a_jU_j. In Thurstonian factor analysis, the consequence is derived by definition of the unique parts of the test and in image analysis by the definition of the common part of the test. The difference is important and involves paradoxes of infinity.

Image analysis proceeds from a multiple correlation approach rather than the partial correlation approach of the Spearman-Thurstone formulation. The image of a test is the least-squares predicted value of the test on the remaining n - 1 tests. In this way a test is broken into two parts: that which is predictable (the image) and that which is not predictable from the other n - 1 tests (the anti-image). The word "image" is used because geometrically it is projection of the test on the space determined by the n - 1 observed tests. (Kaiser, p. 4). The image of a test defines the common part of the test. In factor analysis, the common part of a test is the "image of the test on the space determined by the n - 1 tests in the domain." Images in factor analysis are therefore hypothetical, random variables while in image analysis they are concrete

and clearly defined. (Kaiser, p. 4-5).

Furthermore, Guttman demonstrates that the inverse of the observed matrix of correlations, R, must equal the inverse of an infinite matrix, R, for the conditions of factor analysis to hold. This infinite matrix, representing the domain of tests, must be a diagonal matrix--that is, it must have zeros as off-diagonal elements. Therefore, the inverse of the infinite matrix cannot be the inverse of the observed matrix, (Guttman, 1953, p. 281) and a necessary condition of factor analysis is unattainable.

Guttman then proceeds to show how the breaking of a test into common and alien parts through multiple correlation leads to the derivation of observable common factors. Least squares prediction of a variable, x_{ji}, breaks it into two orthogonal components ρ_{ji} (the image) and e_{ji} (the anti-image or error of prediction). Guttman then demonstrates that though the e_{ji} are orthogonal to a total predictor, x_k, they are not necessarily orthogonal to parts of that predictor. That is, the e_{ji} may be and usually are correlated with the e_k and ρ_k. The Spearman-Thurstone approach, outlined above specifies that these linkages be zero, and, after partialling out common factors, that $r_{jk.c}$ be zero.

Guttman, on the other hand, resolves the observed correlations as follows:

$$r_{jk} = g_{jk} - \gamma_{jk},$$

where:

$$g_{jk} = \sum_i^E \rho_{ji}\rho_{ki},$$

$$\gamma_{jk} = \sum_i^E e_{ji}\, e_{ki},$$

That is, the correlation between two observed variables is equal to the difference between the covariance of their respective images and the

covariance of their respective anti-images. After further resolving the γ_{jk} in terms of the partial correlation of x_j and x_k, Guttman concludes as follows:

> An observed total correlation r_{jk} can be regarded as arising from two sources: (a) the covariance between the common parts of the two variables, and (b) a special pairwise linkage that may remain between the two variables after the remaining n - 2 variables are partialed out.
>
> The Spearman-Thurstone common-factor theory is a special--indeed, degenerate type which specifies zero pairwise linkages. (Guttman, 1953, pp. 288-290).

Image analysis removes the pairwise linkage from the observed correlations and factor analyzes, with the usual procedures, the image covariance matrix of the common parts of the observed variables.

In matrix terms, image analysis may be expressed as follows: (Kaiser)

$$P = R + Q - 2S^2$$

$$Q = S^2 R^{-1} S^2$$

Where:

R = the observed correlation matrix,

R^{-1} = the inverse of the observed correlation matrix,

$S^2 = [\text{diag } R^{-1}]^{-1}$ i.e., the inverse of the diagonal elements of R^{-1},

Q = the matrix of anti-image covariances,

P = the image covariance matrix to be factored in the usual way.

The matrix P has the characteristic of maintaining the Gramian properties of the elements to be factored. That is, the factors extracted all have positive roots and are therefore real not imaginary factors. (Harman, p. 28). The maintenance of the Gramian properties of the matrix is a

distinct advantage of image analysis over multiple-factor analysis in which they are invariably destroyed.

Another advantage of the matrix P is that it generates squared multiple correlations as estimates of communalities. Squared multiple correlations are the well-known best estimate of communalities. They are, in fact, lower bounds of communalities. (Harman, p. 89). Combined with the elements of the image covariance matrix, these "norms" (the Guttman term for the squared multiple correlations) closely approximate the communalities. Using the description "partial" for observed, finite values of images and norms, and the term "total" for their infinite or true values, Guttman summarizes the communality problem as follows:

> Empirically, it has often been found that multiple correlations tend to taper off rapidly as the number of predictors increases. It may often be expected in practice, then, that with n greater than 10 or 15, say, partial images and norms will be so close to their total images and norms that the differences will be negligible, and n can be regarded as virtually "infinite."
>
> On the other hand, it is well-known that the sampling errors associated with multiple regressions can be quite enormous if the sample of people is small. Only large samples of people can be reliably used if n is substantial, and this is as it ought to be. (Guttman, 1953, p. 295).

The n referred to in the above quotation is the number of tests in the inter-correlation matrix. In the study being reported here, the number of tests is 13 and the number of people is 2,069 which conditions should produce close approximations to total images and norms.

The Derivation of Factor Scores

In the basic factor analysis equation given on p. 49, a test is resolved into components. The a_{jm} are test loadings on factors and may be considered the degree to which a test represents a factor or as the

extent to which a test is "saturated" with a factor. The F_{mi} are factor scores and may be considered as the amount of a factor possessed by an individual. In their theoretical distribution, the factor scores of any one factor have a mean of zero and a standard deviation of one; their correlation with the factor scores of any other factor is zero when the factors themselves are orthogonal.

In standard factor analysis the derivation of factor scores is theoretically impossible. (Kaiser, p. 8). The best that can be done is to approximate them via least-squares. (Harman, pp. 337ff). In image analysis, however, factor scores are determinate.

> If, in factor analysis, the number of common factors is small, it may readily be proved that this approximation to factor scores is, in a certain sense, the "worst" possible--the common-factor space is in outer space, maximally removed from the test-space when the number of common factors is a minimum. The approach suggested in this paper stands in stark contrast to this unpalatable situation in traditional factor analysis. Images are not hypothetical; they are observable--the image space is the same as, or is a subspace of the test-space. Consequently the factor scores based on factors derived from the image space are exactly determinable by a straightforward calculation. (Kaiser, p. 8).

Guttman demonstrates that the derivation of factor loadings alone is not sufficient to make factor scores determinate. Depending on the accuracy of the estimate of communalities, "alternative sets of scores for a given hypothesized factor can exist which yield identical loadings and yet correlate negligibly with each other." (Guttman, 1956, p. 276). If the ratio of the true uniqueness of a test to the estimated uniqueness is .5, the correlation between these alternative sets of factor scores is zero. Even if the ratio is as high as .9, the correlation of these sets of scores is only .80. (Guttman, 1956, p. 276). Since a definition of the uniqueness of a test is essentially the definition of its communality,

an accurate estimate of communality is essential for the determination of factor scores.

Formulas for the derivation of factor scores in image analysis are given by Kaiser. (Kaiser, p. 16).

$$f = (A'A)^{-1} A'Wz$$

or when A is non-singular, as is usually the case,

$$f = A^{-1} Wz$$

where

A = the matrix of image factor loadings,

A' = the transpose of A,

$W = I - S^2 R^{-1}$, or the matrix of weights for predicting test scores from the n - 1 remaining tests,

I = an identity matrix, used so that a test is not used to predict itself,

$S^2 = \left[\text{diag } R^{-1}\right]^{-1}$

R^{-1} = the inverse of the observed correlation matrix.

Scalogram Analysis

Studies of school "atmosphere" clearly indicate that attitudes within school organization are important variables having clear influences on pupil behavior and educational output. (Coleman; Herriott; Gordon, 1957; Mitchell; Reiss and Rhodes). In order to include a measure of pupil attitude or motivation towards education, a scale of motivation was developed. Scalogram analysis, a technique derived by Guttman and fully reported in Stouffer, was used to develop a motivation scale.

The aim of scale analysis is to define a scale such that the

responses of individuals to an item are a function of their responses to all other items. In this way, the scale defines an attitude as a continuum in which individuals may possess more or less of an attitude.

> Scale analysis asks the question of whether or not one can consider the responses to a set of items as ordering individuals along the same continuum. The process begins with the definition of a certain universe of interest or universe of attributes which is defined prior to any attempts at data collection. . . .Once the universe has been defined, the items constructed, and the responses obtained, scale analysis investigates whether or not this is scalable in the sense that only one variable is involved and whether or not individuals can be ranked in accordance with the extent to which the variable is present in each individual. (Miklos, p. 56).

After the definition of the universe of content, the next step in scale analysis is the selection of items to represent the universe. If these items are scalable, it will be found that items range in popularity from those which are endorsed by very few to those which are endorsed by almost all respondents. Moreover, a respondent endorsing the least popular item must also be found to endorse the other more popular items.

Similarly, a respondent who rejects the most popular item, must be found to reject the other less popular items. Between these two extremes, respondents must be found who come on the scale with items of a certain popularity and proceed to endorse all items of increasing popularity from that point on. In this way, responses to one item are functions of responses to the other items.

This requirement for the existence of a scale is very stringent and few perfect scales are found. Errors occur in that an individual may endorse an item of a particular popularity or difficulty, but fail to endorse the next item of lesser difficulty. The ratio of errors to the total number of responses leads to a statistic called the "coefficient of

reproducibility." This statistic, with an unknown sampling distribution, is an indication of the degree to which knowledge of a respondent's response on one item enables his response on another to be predicted accurately. (Stouffer, p. 77). A coefficient of .90 is accepted as evidence of the existence of a scale.

Since perfect or near perfect scales are difficult to construct, a further category of "quasi-scale" is defined for scales whose coefficient of reproducibility falls below the arbitrary .90 but which exhibits other scale attributes such as a range of popularity in items and the absence of certain types of error patterns. When scales are not to be used to predict responses but as variables to be correlated with other external variables, quasi-scales can be of considerable use despite their theoretical inadequacy. (Stouffer, p. 163; Miklos, p. 195).

An attitude universe relevant to the achievement of pupils is suggested by Coleman. His study indicates that student attitudes to personal popularity, sports, teachers, and other students achieving high marks are good predictors of general school achievement. Accordingly, a questionnaire was drawn up involving sixteen questions related to general school motivation. This questionnaire is presented as part of Appendix B. The results of this questionnaire were subjected to scalogram analysis.

The methodology of scale analysis is given in detail in Stouffer and has been recently summarized in Miklos. This methodology will not be presented here with the exception of a discussion of the particular technique used which will be given in the chapter on analysis of data.

The Measurement of Socio-Economic Status

There is little agreement among sociologists about precise definitions of the social status of an individual. However, since the time of Weber, it is generally agreed that social stratification in society is complex and multi-dimensional. Weber identified three dimensions of social stratification. The first, economic position, refers to the material possessions and wealth of an individual or group; it is the opportunity or "life chance" which individuals and groups have of obtaining goods and services in the market economy. The second dimension of social status refers to the honorific position or reputation achieved. The third dimension, power, refers to the ability of an individual or group to influence "communal action no matter what its content may be." (Gordon, 1958, pp. 13-14).

Since the time of Weber, other definitions of the dimensions of social stratification have been made but their general form is that of the Weber classification; their meaning may be summed up in Weber's phrase, "style of life." Barber lists three dimensions of social stratification in which contemporary investigations of social class have been cast. The first is the reputation dimension--what people say about the status of others or themselves. The second dimension is defined by the association of people with each other. And the third is defined by the symbolic activities or possessions of people. (Barber, pp. 96ff).

One of the main controversies surrounding the concept of social stratification at the present time concerns the problem of the measurement of social status. The extensive studies in the measurement of social class by Warner lean heavily on the assumption that individuals are

implicitly or explicitly aware of social rank and are capable of translating their evaluations of social rank into ratings which can be communicated to an investigator. (Warner, p. 35). Using this method, Warner derived a rating of social stratification known as Evaluated Participation (E.P.). The application of this method is time-consuming and costly as it involves extensive interviews and knowledge of a community. The E.P. scale has also been criticised for the fact that the extensive interviews required in it, and the personal knowledge required by subjects of each other, limits investigations of the E.P. variety to relatively small communities. The question then arises of whether "social status in America" can be described by studying in depth only a few of its smaller communities. (Gordon, 1958, p. 104). To answer these criticisms, Warner developed a further scale, the Index of Status Characteristics (I.S.C.), which relies on the rating of objective characteristics and which can be easily and widely applied. (Warner, p. 39).

Warner describes I. S. C. as an index of socio-economic status which includes four characteristics: occupation, source of income, housing type, and dwelling area. (Warner, p. 40). He claims, further, that this scale is not a measurement of social class, as the E. P. scale is, but an index of it. This statement may well be semantic hair-splitting as he adduces as evidence of the validity of the E. P. scale the fact that it has a multiple correlation of .974 with the four sub-scales of I. S. C. (Warner, p. 174).

In general, then, the measurement of social status may be seen as an attempt to quantify both the inner or qualitative aspects of social status and its outer or objective aspects.

Occupational scales. Occupation is one of the most easily objectified indices of social status. Kahl and Davis confirm the high correlation between occupation and social status reported by Warner and derive a social status factor with highest loadings on occupational variables.

An objective occupational scale devised for use in Canada is the Blishen Occupational Scale. (Blishen, p. 519). This scale, based on Canadian census data, yields a standard score for each of 343 occupations weighted equally on mean income for the occupation and years of schooling required to attain it. Scores on the scale may be regarded as T-scores since they have a mean of 50 and a standard deviation of 10 with a range from 32.0 to 90.0. The highest scores are given to judges, dentists, physicians, and surgeons; the lowest are given to bootblacks, fish canners, hunters and trappers. Mean scores are given to occupations such as those of policemen, office clerks, farmers, and mechanics.

The Home Index Scale. An example of a scale which measures more subjective aspects of social status is the Gough Home Index Scale. (Gough). The scale is administered directly to the subject who then makes a self-report by answering twenty-one questions of the yes-no type. It is designed with readability standards appropriate to junior high school students. A re-test reliability of .989 is reported for it and a Kuder-Richardson reliability of .74 is given for high school students. A modification of this scale, as made by Elley, was adopted for use in this study. This scale is given in full in the Pupil's Questionnaire to be found in Appendix B.

The Measurement of Educational Ability

Ability has been defined as "a response subject to voluntary control." (Cronbach, p. 29). A response is thus observed behavior whatever its source or "cause" may be. This definition avoids an argument, which is not relevant here, as to whether intelligence as an ability is innate or learned capacity. Tests of mental ability are those which seek to measure abilities "which are valuable in almost any type of thinking or learning." (Cronbach, p. 30). Tests of mental ability do, however, assume that the response measured is a maximum performance.

The Spearman approach to the measurement of intelligence hypothesized the existence of a single general factor, g, which would account for all the intercorrelations between tests of academic achievement. In this theory, the amount of g possessed by an individual would determine his performance on all major mental acitivites such as cognition, problem solving, and learning. In fact, virtually all human performance would be dependent upon this single factor; g is thus "innate and ineducable." (Vernon, 1950, p. 13). Spearman's theory has since been modified by the development of multiple-factor analysis and the concept of group factors. (Thurstone). A group factor is one on which certain tests have loadings but others have none. Or, as Harman states it, a group factor is one that is "present in more than one but not in all variables." (Harman, p. 12). Without the concept of group factors, all tests are measures of g. Educational and vocational guidance on the basis of tests results are therefore impossible; in fact, the only tests that are useful are those yielding the purest measures of g. (Vernon, 1951, p. 15).

The group factor theory of intelligence holds that there are certain

types of ability over and above g. Group factor theories then do not dispense with a factor g but add other factors to it. The English school of factor analysists identifies two further factors--the v:ed and k factors. The k factor is a measure of mechanical ability and the v:ed factor is a measure of verbal-educational achievement. (Vernon, 1956, p. 144). The v:ed factor is of considerable importance for this study as it suggests a criterion for the evaluation of school systems. As a measure of learned ability free of the more basic g it may be used as a measure of educational output. Two recent studies of achievement in samples of Alberta school children have identified a v:ed factor. (Elley; Hukins).

> The second factor has its highest loadings on two vocabulary tests, but since it is also heavily weighted in the arithmetic tests, it appeared to be the familiar v:ed factor, which recurs so often in English analyses of school children. (Elley, p. 116).

The Elley study also identifies a g factor with heaviest loadings on three culture-reduced tests--the Raven Matrices, the Cattell Test, and the Figure-Analogies sub-test of the Lorge-Thorndike battery. These are all tests intended as measures of intelligence free from cultural and educational learning. In this study, the g factor accounted for 51 per cent of common variance and the v:ed factor for 25 per cent.

The study by Hukins used, in addition to measures of achievement in Science, the Cooperative School and College Tests of educational proficiency (SCAT). Hukins identifies the v:ed factor derived in the study as follows:

> Factor I is interpreted as a verbal-educational factor. . . .All the tests of achievement in science have high loadings on the factor except the Test of Understanding Science. . . .It is apparently not a specialized science ability, however, because the verbal tests also have high loadings upon it. The moderate loadings of SCAT Quantitative and the general reasoning test also lend weight to this interpretation.

The v:ed factor found in this study accounted for 31 per cent of common variance.

The findings of these studies indicate that achievement tests, like those of the Alberta Departmental Examinations, can be resolved into factors representing general intelligence and educational attainment. It may be presumed, too, that the manner of operation or the steady state within educational systems will have a decided influence upon the educational achievement factor.

CHAPTER IV

THE COLLECTION OF DATA

The collection of data for the study was begun in May, 1962 and completed in the fall of the same year. Clearance for the study was gained from the Alberta Department of Education and from the Alberta Teachers' Association. Twenty-five school districts in the province were large enough to fit the design of the study. There are many more than twenty-five districts in the province, but most of them had less than the required two schools or else had fewer than the required two Grade IX classes in each school. Letters were then sent to the superintendents of the twenty-five districts requesting permission to carry out the study in two schools of their district. Of these superintendents, twenty-three gave permission for the study to be undertaken in their districts.

After securing permission from the superintendent, two schools within the district were selected randomly where there was a choice among a number of schools. The principals of these schools were then contacted and their permission to carry out the study in their particular schools was requested. All principals replied and all but one gave permission for the study. Thus out of twenty-five districts originally contacted, twenty-two of them were available as subjects for the study. This sample produced the forty-four schools and eighty-eight classrooms in the way required for the analysis of variance design. Where there was more than one Grade IX classroom in a school, two classes of the total were again

randomly assigned.

Data from Questionnaires

Two sets of questionnaires were then sent to the school involved. One questionnaire was completed by the principal. This questionnaire concerned the organizational characteristics of the school and personal characteristics of himself. The second set of questionnaires was completed by the Grade IX teahcers of the classes in the sample. This questionnaire involved the educational background, experience, and other personal characteristics of the teacher. The form of these questionnaires is based on a similar questionnaire devised in connection with research done for the Alberta Royal Commission on Education. (MacArthur and Linstedt). The two questionnaires sent to principals and teachers are presented in Appendix B.

Each classroom was then visited by the author and a third questionnaire administered to the pupils in the classes. This questionnaire contained questions related to pupil attitudes and activities, two scales of socio-economic status and a scale composed of sixteen questions designed to measure educational motivation. It was this last scale which was subjected to scalogram analysis as described in the previous chapter.

The frequency distribution of responses to these three questionnaires is given in Appendix B. This information was also recorded on I.B.M. punch cards.

Achievement Data

Data on the achievement of pupils in the classes were obtained from records of the Alberta Department of Education. Each Grade IX pupil

in Alberta schools is required to write a series of examinations of subject matter knowledge based on material from the Alberta Grade IX curriculum. The subjects covered are English Language, English Literature, Social Studies, Science, and Mathematics. At the same time that these examinations are written, pupils also write the Cooperative School and College Ability Tests (SCAT) and the Sequential Test of Educational Progress Reading (STEP).

The results of all these tests for each pupil in the sample were obtained from Departmental records in I.B.M. punch card form. Each card also contained a Departmental pupil identification number. In filling out the pupil questionnaires, each pupil was asked to place his name on the copy. By matching pupil names with their Departmental code numbers, it was possible to produce an amalgamated card deck combining information about pupil achievement with other information such as socio-economic status of the pupil.

The Sample

Complete data were available on 2,069 students. Approximately 150 pupils were discarded from the sample because of incomplete data. Most of these were accounted for by pupils who withdrew from school between the time the pupil questionnaires were completed in May and the time the Departmental Examinations were written at the end of June. The rest were discarded because of the lack of complete socio-economic data. Other pupils missing from the sample were those who were absent from school on the day the investigator visited their classrooms.

The districts involved in the sample represented both rural and urban areas. The districts may be divided into three groups: cities,

counties, and divisions. The city districts are made up of two large and four smaller urban areas in the province. Counties and divisions in Alberta school government are units of educational administration in non-urban areas of the province. Pupils in counties and divisions are drawn from town and rural areas. Of the 2,069 pupils in the total sample, 624 are from city districts, 516 from county districts, and 944 from divisions. Mean achievement scores on the five Departmental Examinations of subject matter knowledge are all close to 60 which is approximately the mean imposed by Departmental scaling on pupil marks. Altogether there were approximately 20,000 pupils who wrote the Grade IX Departmental Examinations in 1962. This sample, though not selected by rigorous stratified, random sampling may be considered a representative sample because of its size and its range of district, school, and classroom characteristics. Virtually all of the districts fitting the requirements of the analysis of variance model were included in the sample. Only three of the possibly twenty-five were not represented. In larger districts, schools and classes were selected from the several which were available. However, in many districts, the schools and classes selected represent a large, if not total, proportion of those available. The sample is thus close to the maximum possible within the Alberta schools system, given the requirement that there be two schoos per district and two classes per school.

Distributions of the achievement scores of the sample are given in Appendix C. The distributions presented there resemble the provincial distributions. (MacArthur and Hunka). The characteristics of schools, principals, and teachers in the sample, as shown in Appendix B, also have as wide a variability as might be expected in the province as a whole. (MacArthur and Linstedt).

CHAPTER V

THE TEST DATA

This study involves the analysis of thirteen tests of academic achievement, socio-economic status, educational ability, and motivation. For each test, scores were obtained for the 2,069 pupils in the sample. These tests are described in this chapter. The analysis of the large number of data involved in the study relies heavily upon the use of a computer. The computer used was the I.B.M. 1620 at the University of Alberta. The number and nature of the programs used are also described in this chapter.

Tests Used in the Study

Of the thirteen tests used in the study, three are tests of socio-economic status, two are tests of educational motivation, five are tests of academic achievement, and three are tests of educational ability.

The Home Index Scale. The Home Index scale is a modification of the questionnaire constructed by Gough. This modification, made by Elley, is composed of fourteen items from the original scale and five new ones. All items were tried in a pilot study and the twenty best items, as indicated by item analysis, were retained. Items deleted from the original scale were judged to be inappropriate in Alberta, such as an item indicating the possession of a heating system in the child's home. The reliability coefficient of the modified scale, calculated by the corrected split-half method was .77. (Elley, p. 102). The Gough

scale is discussed in more detail in Chapter III in the section on the measurement of socio-economic status.

Each pupil completed a home index scale at the time the Pupil Questionnaire was administered in each classroom. Scoring of the test is accomplished by the simple addition of the number of items to which the respondent has replied "yes". The maximum possible score is thus twenty and the minimum zero. The distribution of responses to this scale is presented in Appendix C.

The Blishen Occupational Scale. As reported in Chapter III, this scale represents Canadian occupations as T-scores based on an equal weighting of the income and educational training associated with the occupation. The full range of scores to be assigned to occupations is given in Blishen and in Elley. Also of assistance in identifying occupations on the scale is the Classification of Occupations (Canada: Bureau of Statistics) which gives a full range of occupational titles in Canada and assists in relating such titles to occupations appearing on the Blishen scale. Each pupil was asked to identify, on his copy of the Pupil Questionnaire, the occupation of his father and his mother. These occupations were then assigned scores according to the Blishen scale. Where there was no father or where the father was not working, the occupation of the mother (if gainfully employed) was used. The distribution of scores assigned for this scale is presented in Appendix C.

In assigning scores to the parental occupations of rural children, a weakness of occupational rating scales became apparent. The parents of many rural children are farmers and the Blishen Occupational Index assigns

only one score to farmers--49.2. No distinction, thus, is made between kinds of farmers or the size of farm operated, or whether they own or rent the farm. This lack of differentiation caused a considerable kurtosis of the occupational distribution.

The Socio-Economic Status Scale. The Blishen Occupational Scale and the Home Index Scale represent two aspects of social status. The Occupational Scale may be thought of as representing one of the external aspects--the position of an individual in a stratified social system rather than the way his life is lived. The Home Index scale may be thought of as representing more qualitative aspects of social position--the number of books in the home, high school and college education, family vacationing, etc. In order to obtain a scale combining both these aspects, scores on the Home Index Scale were converted to T-scores and added to the scores obtained on the Blishen Occupational Scale.

For example, it was observed that children of farmers who would be assigned the single score 49.2 on the Occupational Scale varied in Home Index scores from the highest to the lowest. A combination of scores, then, might possibly give a more accurate description of social status. The distribution of these combined scores, presented in Appendix C shows markedly less kurtosis than the distribution of Occupational scores.

Study hours. One of the questions on the Pupil Questionnaire asked students to give a self-report on their study hours as follows:

> 5. HOW MUCH TIME, on the average, do you spend on HOMEWORK outside of school hours?

CHECK ONE

1. . . . .None or almost none
2 . . . .About ½ hour per day
3 . . . .About 1 hour per day
4 . . . .About 1½ hours per day
5 . . . .About 2 or more hours per day

Responses to this question were accepted as an indication of motivation and were included in the test battery for intercorrelations. The distribution of responses to this question are included in Appendix B with the Pupil Questionnaire.

Motivation. A pupil score on a scale of motivation was obtained by the scalogram analysis of sixteen questions selected to represent the universe of the attitude "educational motivation." A discussion of the scalogram analysis is presented in one of the following sections of this chapter. Only eleven of the original sixteen questions were found to scale reasonably. These eleven questions, however, gave a coefficient of .83 which is below the arbitrary standard for a scale of .90. However, these questions were accepted as forming a quasi-scale. Scores are formed by counting each question answered positively; scores are thus out of eleven possible. The distribution of this scale is given in Appendix B.

The Alberta Departmental Examinations. The Alberta Departmental Examinations of subject matter knowledge may be considered as standard achievement tests administered under standard conditions. They are not, however, quite the same type of test as is included in the term "standardized tests of achievement." Standardized tests of achievement are proficiency tests normed to represent achievement in particular grades and subjects. The norming is typically done on large groups of students in different

school systems and in widely differing geographical locations. Standardized achievement tests are usually of unquestioned reliability since they are subjected to careful analysis and modification in the process of construction. They are, however, of dubious validity in some cases. (Cronbach, p. 364). The type of validity problem which arises is that of content validity. The establishment of content validity is largely a logical problem--does this test measure the content which I want to measure? The establishment of the content validity is thus largely up to a school, or system, or to whomever is using the test.

The Alberta Departmental Examinations are of unknown reliability; but their content validity is clearly established since they are the sanctioned yardstick for the measurement of important aims achieved in Alberta Grade IX education. Many vocational and educational choices for students are also made on the basis of their achievement on these examinations. As indicators of freshman success in university, the Grade IX examinations show a significant predictive validity. (Black, p. 229).

The marks obtained by pupils on the Examinations are scaled according to a set of examinee standards. The scaled scores are out of one hundred, in most cases, and are assigned as follows:

80 - 100	Top 10 per cent of the pupils
65 - 79	next 25 per cent of the pupils
50 - 64	next 35 per cent of the pupils
40 - 47	next 20 per cent of the pupils
0 - 37	bottom 10 per cent of the pupils. (MacArthur and Hunka, p. 11).

One examination, English Literature, is scaled from zero to fifty.

The examinations are prepared and conducted each year under the supervision of High School Entrance Examination Board. Papers are set in

English Literature, English Language, Social Studies, Mathematics and Science with different papers being set each year. Questions on the papers are largely objectively scored questions, although some papers have short answer and essay questions as well.

In the case of the Social Studies and Science examinations, the scaled test scores are combined with a scaled score derived from teachers' estimates of pupil achievement. The scores used in this study are those from the examinations only and do not include the teachers' estimates. Distributions of achievement scores are given in Appendix C.

In view of their scaling, size of sample, and content and predictive validity, scores from these examinations may be considered consistent as well as valid.

The Sequential Test of Educational Progress: Reading (STEP: Reading). The STEP test of Reading is administered to all Grade IX Alberta students during the same period in which they are writing Departmental Examinations. This test calls for "ability to understand direct statements, to interpret and summarize passages, to see motives of authors, to observe organization of ideas, and to criticize passages with respect to ideas and purposes of presentation." (Buros, p. 64). The test is a typical standardized achievement test with the usual care and technical competence in evidence in its construction. (Buros, pp. 750-754). A Kuder-Richardson reliability coefficient of .90 is reported for the test. No information on predictive or concurrent reliability is given, but with reading tests of this kind content validity is the essential component. The reviewers in Buros considered this test to have good content validity.

The Cooperative School and College Ability Tests (SCAT). The SCAT battery of tests is described by Cronbach as lying mid-way between tests with a maximum educational loading and those with a minimum. (Cronbach, pp. 234-237). Thus SCAT falls between tests that are measures of subject-matter proficiency and tests that are designed primarily to measure more pure ability such as Spearman's g. The SCAT manual emphasizes its role as a predictor of academic success.

> The tests in the SCAT series have been designed and developed for the principal purpose of helping teachers and counselors--and students themselves--to estimate the capacity of each individual student to undertake the academic work of the next higher level of schooling. . . . (Educational Testing Service).

SCAT is the product of technically expert test construction and reviews of it are cautiosly optimistic about it.

> Undoubtedly, SCAT is a superior test series. It clearly shows the result of careful planning, an excellent experimental programme, and the use of sound, up-to-date statistical procedures. (Buros, p. 455).

In studies of concurrent validity, correlations ranging from .50 to .70 were obtained between the various scales of the test and school grades. (Buros, p. 451).

The SCAT test battery is composed of four sub-tests from which three scores may be derived--a verbal, a quantitative, and a total score. The meaning of these scores is perhaps the weakest point in the tests.

> There are a number of. . .deficiencies in the SCAT. In the first place, a number of doubts are raised by the quantitative score: in the upper grades the items are apparently too easy, at least for superior students. . . .The validity study shows disturbingly high intercorrelations between the quantitative and verbal score. . . . One wonders whether the quantitative subtests are indispensible and whether their omission would greatly reduce the effectiveness of the test. (Buros, p. 455).

Two subtest scores of the SCAT battery are used in this study--the Verbal and the Quantitative--despite the reported lack of independence. This choice was made in an effort to define verbal and quantitative abilities in the tests of subject matter knowledge.

Summary and Enumeration of Tests. For convenience of reference in the remainder of this study, the thirteen tests are named and numbered below. An abbreviation in parentheses follows the full name of the test. These abbreviations will be used in referring to the tests in tables to be found in following sections of this report. The tests are numbered in the order in which they are found in the intercorrelation matrices and factor analysis matrices.

1. Socio-Economic Status. (S.E.S.)
2. Home Index Scale. (Home)
3. Blishen Occupational Scale. (Occup.)
4. Study Hours. (Study)
5. Motivation Scale. (Motiv.)
6. The Sequential Test of Educational Progress: Reading. (Rdg.)
7. Department of Education: English Literature. (Lit.)
8. Department of Education: English Language. (Lang.)
9. Department of Education: Social Studies. (Soc. St.)
10. Department of Education: Mathematics. (Math.)
11. Department of Education: Science. (Science)
12. Cooperative School and College Ability Tests: Verbal. (V score)
13. Cooperative School and College Ability Tests: Quantitative. (Q score)

The Use of Computer Programs

Educational research has been considerably modified by the advent of high-speed computers. (Flanagan). The availability of computer analysis places studies within the range of possibility which previously were too complex or lengthy to handle with hand or desk-calculator analysis. The analysis of data in this study relied heavily upon computer handling of data. The sequence and nautre of the programs used are listed in this section. The computer used in the analysis was the I.B.M. 1620 at the University of Alberta and most of the programs used were supplied by the University of Alberta Computing Center. One of the programs was written by Dr. S. M. Hunka of the Department of Educational Psychology at the University of Alberta and two sequencing programs designed to facilitate the transfer of data from one program to another were written by the author from I.B.M. programing guides. (Leeson and Dimitry).

Intercorrelation program. This program was used to intercorrelate data on the thirteen tests variables for the 2,069 pupils in the study.

Inverse of matrix program. The matrix of intercorrelations on the thirteen variables was inverted for use in the image analysis formulas:

$$Q = S^2 R^{-1} S^2$$

and $$W = I - S^2 R^{-1}.$$

For discussion of these formulas, see Chapter III.

Eigenvalue-Eigenvector program. The method of principal axes factoring is a mathematically precise and psychologically meaningful method of factor analysis. (Harman, pp. 154ff). The derivation of

principal factors involves the solution of entitites known as eigenvalues and eigenvectors. (See the discussion of these in the section on analysis of data). Although the method of principal axes factoring has been known since the 1930's, solution of them by hand calculation is so laborious that the derivation of principal factors is practical only through the use of high-speed computers.

The matrix P from the image analysis was factored into principal components using a program based on the Householder method. The matrix P was resolved into thirteen eigenvalues and eigenvectors in approximately twenty minutes by the I. B. M. 1620.

Varimax rotation program. A varimax rotation was used to rotate the principal axes factors to a "simple structure" position. The aim of simple structure is to aid in the interpretation of factors. The varimax rotated factors are mathematically equivalent to, but "psychologically" more meaningful than the derived principal axes factors. (Harmon, pp. 289ff). This technique is discussed more fully in the following chapter on the analysis of data.

The principal axes factors derived from the eigenvalue-eigenvector program were then repunched for the Varimax rotation program. The rotated factors were then used in the derivation of factor scores. That is, the matrix of rotated factor loadings, A, was inverted to A^{-1} and used in the factor score formula which appears below.

Standard score program. This program was used to change the test data from the 2,069 pupils on thirteen tests into standard score form.[1]

---

[1]This program was written by Dr. S. M. Hunka.

Data in standard score form is required to resolve test data into factor scores according to the following image analysis formula:

$$f = A^{-1} Wz,$$

where z is the matrix of standard scores of size 13 X 2069.
In order to make the output of the standard score program conformable to the input of the factor score program, the standard score program was modified by the author. In this way, no repunching of the very large z matrix was required.

Matrix multiplication. A matrix multiplication program was used to assist in forming the matrix P required for image analysis.

The derivation of factor scores is also essentially a multiplication of matrices, and this program was again used to form the factor scores. The A^{-1} matrix given in the formula immediately above was obtained from the matrix inversion program.

Data conversion program. The large output from the matrix multiplication used in the derivation of factor scores was not in a format or numerical form conformable to the analysis of variance program. Therefore another program was written to change the data into the required numerical form and to arrange the format and coding needed for the analysis of variance program.

Analysis of variance program. The analysis of variance program was used on two sets of achievement data. Analysis was first performed on the scaled score results of the 2,069 pupils on the Department of Education Examinations. The achievement factor scores were also analyzed with this

program. By coding all data according to class, school, and district, this program yields all the required sums of squares for the analysis of variance model used in this study.

Card Punching and Handling

One of the major practical problems in the study was the handling of data--there were over 25,000 pupil scores on the thirteen basic tests. To facilitate handling and to reduce the chance of error, the repunching of data cards was kept to a minimum. The largest punching task was the punching of replies from the Pupil Questionnaire. Once this punching was done, cards in the Pupil Questionnaire deck were matched card for card with a deck obtained from the Department of Education containing the achievement, STEP, and SCAT data. This matching was done visually by comparing code numbers on the two decks. To help in this process, both the Pupil Questionnaire deck and the Department of Education deck were "interpreted." That is, the data and code numbers were spaced and printed in easily readable form at the top of the cards. The two decks were then amalgamated using the I. B. M. 514, which is a card reproduction unit. A single deck was thus produced with the data from the thirteen variables arranged in a format conformable for input to the intercorrelation program, the standard score program, and the analysis of variance program. The data conversion program, described above was used to transfer data from the factor score program to the analysis of variance program. In this way, data on the thirteen variables required only the one punching, despite their use in a variety of programs.

The punching of data for matrix multiplication, image analysis, and

matrix inversion was relatively easily done since the matrices were small and easily checked for error.

CHAPTER VI

THE ANALYSIS OF DATA

Specific techniques used in the analysis of variance, image analysis, rotation of axes, scalogram analysis, and derivation of factor scores are discussed in this chapter.

Analysis of Variance

The analysis of variance data were arranged on I. B. M. cards as follows: data field test #1, data field test #2, data field test #n, class code, school code, district code. The analysis of variance program allows various sorts and calculations in a single pass of the data. For example, a data field for a test may first be sorted and analyzed on the class code, then the school code, and the district code. In each of these sorts, a one-way analysis of variance is performed with the data constant but with the classification into groups changing according to the code number on which the sort occurs. Each sort of a single data field on each code field produces a different estimate of within and between variance although the total is constant. That is, one-way analyses of variance are performed first over classes, then over schools, and finally over districts. In this way, total variation is broken into sources from classes, schools, districts, and residual effects, as stated in the analysis of variance model in Appendix A. The computation formulas also given in Appendix A will serve to make the derivation of the sources of variation more clear.[1]

---

[1]The author is indebted to Mrs. P. Montgomery of the University of Alberta Computing Center for deriving these computing formulas.

In summary, however, it may be said that the one-way analysis of variance over classes produces estimates of the within (residual) variation and of the total between variation. The total between variation (sums of squares) is made up of sources from classes, schools, and districts. Then the one-way analysis of variance over schools gives the between variation due to schools and districts alone. Subtracting these two sources of between variation, gives the variation due to classes alone. Then the one-way analysis of variance over districts yields the variation due to districts alone. Subtracting this variation from that due to schools and districts gives variation due to schools alone and to districts alone. The three sources of between variation calculated in this way sum to the total between variation calculated in the one-way analysis of variance over classes.

Assumptions of analysis of variance. Three important assumptions made in analysis of variance are normality of errors in the population and in models, equality of variance in the errors, and the statistical independence of the errors. (Scheffé, p. 331). The effects of departures from these assumptions have been investigated but complete knowledge of these effects has not been obtained, as Scheffé states, because assumptions may be violated in more ways than they can be satisfied.

However, it is recognized that the assumptions underlying statistical procedures are often apt to be violated in practical applications. Some assumptions, indeed, are introduced "mainly to ease the mathematics of the derivation." (Scheffé, p. 360). There are cases, then, in which the assumptions of a statistical model may be violated in a practical

application without invalidating conclusions derived by using the model, despite the departure from its assumptions. It is also possible to develop tests appropriate to the model when its assumptions have been altered or violated. However, in the case of analysis of variance, such tests though possible in theory "would be hopelessly complicated to carry out in any but the very simplest cases." (Scheffé, p. 362).

Scheffé is cautious in advocating departures from the assumptions of analysis of variance but regards difficulties stemming from lack of independence "as the most formidable to cope with." (Scheffé, p. 364). Authors such as Winer and Edwards accept the conclusions of the Box studies which indicate that the F-test is "robust" under violations of the assumptions of normality and equality of variance. Such effects as there are tend to increase the probability of type-I errors. Effects upon Type-II errors are less clear. (Scheffé, p. 357). Thus Winer concludes in regard to the F-test that:

> The magnitude of type 1 error is not seriously affected if the distributions depart moderately from normality or if the population variances depart moderately from equality; i.e., the test is robust with respect to the assumptions of normality of distribution and homogeneity of error variance. (Winer, p. 61).

With regard to lack of independence in observations, the most serious error is introduced by serial correlation--that is, by a sequential influence from one observation to another. Plotting rainfall by hours of the day introduces a serial correlation in that rain in one hour of the day is likely to mean rain in the next. (Scheffé, p. 359). Another type of dependence is introduced into observations by scaling them or standardizing them since independent scores are made to fit a particular distribution.

In regard to the data of this study, tests of their appropriateness to the assumptions of analysis of variance would be difficult to make in view of the variety and size of the data involved; even Scheffé does not recommend that such tests always be made before performing the analysis of variance. (Scheffé, p. 83). Therefore an assumption of the study is that the data involved do not go beyond the "moderate" departures from the assumptions of normality and equality of variance accepted by Winer and Edwards. As to a lack of independence within the data, the analysis of the factor scores and scaled achievement scores may be open to objection since scaling and standardizing introduces an internal correlation in data. However, in this case, the size of the sample will reduce such correlation to a very small value. Therefore, it was decided to accept the data of the study as approximating the assumptions of analysis of variance closely enough to permit the analysis in this case.

Image Analysis

The procedures of image analysis involve the solution for P in the equation,

$$P = R + Q - 2S^2$$

R is the matrix of intercorrelations on the thirteen test variables and Q is the "anti-image covariance matrix" or the matrix of non-common parts of the test intercorrelations. Adding this matrix, Q, to R removes the non-common portion of the intercorrelations and leaves the common portion. However, the diagonal elements of R + Q are equal to one plus the anti-image of a test; that is, one plus the unique or non-common part of a test. To modify these diagonal elements, diagonal matrix $2S^2$ is subtracted.

The matrix $2S^2$ contains double the value of the anti-images in its idiagonal and zeros in its off-diagonal. Therefore, subtracting it from R + Q leaves the Gramian matrix P of image covariances. P then contains the common portions of test intercorrelations and squared multiple correlations in its diagonal. The matrix P is then factor analyzed according to whichever factor analytic method is considered appropriate.

Principal axes factoring. Principal axes factoring is the most satisfactory technique for factor analysis in that it is mathematically precise and psychologically meaningful. The resolution of a correlation matrix into principal axes factors requires the solution of the equation,

$$Rx = \lambda x$$

where:

R = the matrix to be factored

x = the eigenvector

λ = the eigenvalue.

The equation above is known as the "characteristic equation," and may be read as: the matrix of intercorrelations times a scaling vector equals the characteristic root or eigenvalue times the same scaling vector. The derivation of factors from the equation requires solution for λ and for x.

Values for are obtained from the equation which states that the determinant of $R - \lambda I$ equals zero. Substitution of these lambdas or eigenvalues in the principal axes equation above yields the x's or eigenvectors. Each eigenvector has an associated eigenvalue; rescaling the elements of the eigenvectors by the square root of their associated lambdas

yields the principal axes factors. (Harman, pp. 154ff).

There are as many factors in the correlation matrix as there are solutions for lambda in the determinantal equation given above, and each lambda yields its associated eigenvector.

The meaning of principal axes factors. The resolution of an intercorrelation matrix into principal axes factors requires that the first axis of a factor be placed so that it accounts for a maximum of variance in the tests of the intercorrelation matrix. The second factor, placed at right angles to the first, accounts for a maximum of the remaining variance, and so on until all test variance is accounted for. The variance of a factor is the sum of the squares of test loadings on it. This variance of a factor is precisely the value of the lambda. Thus the knowledge of lambda values indicates the size and importance of the factors associated with them. The tests having the highest loadings on the factor serve to identify and name the factor.

In the present study, the image matrix P was factor analyzed using the principal axes method.

The Rotation of Axes

Given an initial solution of factor space for a battery of tests, the initial factors may be transformed or rotated to an infinite number of other positions that still satisfy the basic factor analytic equations outlined in Chapter III. As Guttman puts it, once the original factors have been determined, mathematics has "nothing to say" as to which rotation is preferable to another. (Guttman, 1954, p. 256). Various criteria for the rotation of axes have been suggested, but their aims are usually

similar to the criterion known as "simple structure," which was suggested by Thurstone. (Thurstone, p. 338). Simple structure is an attempt to represent factors in the most parsimonious manner possible in the factor space defined. For the simple structure criterion to be met, test loadings on a factor must tend towards zero or unity. In this way, factors may be interpreted "simply" in terms of a few tests rather than many.

The varimax rotation, developed by Kaiser, is one of the most useful methods of rotation. (Eyman, et al.). It is an analytical method--as opposed to intuitive hand methods--which avoids the indeterminacies of hand methods and is described as follows:

> . . .The simplicity of factor p is defined as the variance of its squared loadings. . . .when the variance is at a maximum, the factor has the greatest interpretability or simplicity in the sense that its components. . .tend toward unity and zero. The criterion of maximum simplicity of a complete factor matrix is defined as the maximization of the sum of these simplicities of the individual factors. . . . (Harman, p. 301).

The varimax criterion has the advantage of not weighting one test over another in the rotation; all contribute equally to the criterion to be maximized. Moreover, it avoids the single general factor which other analytical rotation criteria often produce. Its strongest feature, however, is that it tends towards factorially invariant solutions. That is, the psychological description of a factor remains constant over changes in the composition of the test battery. (Harman, p. 308).

The principal axes factors derived in the present study were rotated using the varimax criterion.

The Derivation of Factor Scores

In factor analysis, the formula for the derivation of factor scores

is

$$F = S' R^{-1} Z,$$

where:

F = the matrix of factor scores,

S' = the transpose of the matrix of structure loadings,

R^{-1} = the inverse of the correlation matrix of tests,

Z = the matrix of test scores in standard score form.

In image analysis this formula is replaced by the following:

$$f = A^{-1} W z,$$

where:

f = the matrix of factor scores,

A^{-1} = the inverse of the matrix of test loadings on factors,

W = the matrix of weights for predicting a test score from the n - 1 other tests in the battery,

z = the matrix of test scores in standard score form.

The multiplication of A^{-1} by W produces a matrix of weights which when multiplied by the matrix of standard scores produces the factor scores. Thus each test is assigned a weight for each factor. Weights may be positive or negative. The factor score of individual i on factor m is the sum of the product of the n test weights for factor m and the N standard scores for individual i, or:

$$f_{mi} = \sum_{j}^{n} W_{mj} Z_{ji}$$

In this way, a factor score of an individual depends on the weights assigned tests in a factor and upon the performance of the individual on the battery of tests.

The theoretical distribution of factor scores for a given factor has a mean of zero and a standard deviation of one.

In the present study, factor scores for the four largest factors were obtained. The dimensions of the matrices in the factor score formula were:

$$4^{f}2069 = 13^{A^{-1}}13^{W}13^{Z}2069$$

The matrix A in this formula is the matrix of rotated factor loadings. That is, the matrix A of principal factor loadings were rotated by the varimax criterion and it is the rotated matrix A which was inverted and used in the above equation.

The analysis of factor scores. As stated in Chapter III, the factor score of an individual is a measure of the amount of the factor possessed by the individual. One of the hypotheses of this study states that the battery of tests of achievement, motivation, socio-economic status, and educational ability may be factored to obtain a factor representing achievement, the typical v:ed factor, on which the achievement tests have high loadings.

After extracting and identifying this factor, a further aim of this study is to subject the factor scores of the achievement factor to the analysis of variance in the same way that the scaled achievement scores were analyzed. In this way, it may be presumed that these factor scores will reduce residual variance (pupil differences) and increase the system effects of class, school, and district.

Scalogram Analysis

One method of scalogram analysis, as described in Stouffer and

applied in Miklos, involves the construction of a scalogram board. This board consists of one hundred slats each of which has one hundred holes drilled in it at equal intervals. The slats are placed horizontally across the board and represent the responses of individuals to the items being scaled. The item responses are represented by columns. The first column is chosen to represent the least popular or most "difficult" item in the attitude scale. That is, the first column represents responses to the item endorsed by fewest individuals. The next column represents the next most popular item, and so on through to the item endorsed by the greatest number of individuals.

On a dichotomous scale, endorsation would be construed as answering, say, "yes" to a yes-no question. If this scale consisted of six questions (the number of items which can be found scalable is usually small), the first six columns would represent the "yes" answers to the questions in ascending order of popularity and the next six columns the "no" answers to the six items in the same order. By moving the rows or slats to higher or lower positions on the board according to the number of positive responses, a visual check may be made for the"parallelogram" of responses which indicates the presence of a scale. Errors are seen as gaps in the parallelogram. That is, errors occur when an individual has endorsed a question of a certain popularity but not the next question of greater popularity. The slats are arranged on the board so as to minimize the number of errors. The coefficient of reproducibility is calculated by counting the number of errors, dividing their sum by the total number of responses, and subtracting the resulting fraction from unity. That is:

$$\text{Rep} = 1 - T/N$$

The scale score of an individual is assigned by counting the number of positive responses. Figure 2 contains a scalogram pattern which might be obtained for a dichotomous scale of six items. Item #1 is the least popular since only one person has endorsed it and item #6 the most popular since all individuals endorsed it. The response of individual "f" contains errors since he endorsed item #2 but not items #3 and #4.

		"Yes" Response	"No" Response
	Scale Score	Item 1 2 3 4 5 6	Item 1 2 3 4 5 6
a	6	x x x x x x	
b	5	x x x x x	x
c	4	x x x x	x x
d	4	x x x x	x x
e	3	x x x	x x x
f	3	x x x	x x x
g	1	x	x x x x x

FIGURE 2

SCALOGRAM OF SIX DICHOTOMOUS ITEMS

Not all scales need be dichotomous but in practice scales with more than two response categories are rarely found to meet the criteria of scalogram analysis.

In the present study, the scale presented to pupils contained sixteen items related to the attitude "motivations towards education." The items required the respondents to indicate their attitudes towards the items on a five point scale. (See Appendix B). The five point scale was chosen in an effort to gain a scale using more than two categories even if the total five did not prove feasible. Placing the responses of a sample of one hundred pupils on the scalogram board soon revealed that the full number of categories was unworkable since that number of categories

failed to make consistent distinctions among pupils. The best patterns were obtained by regarding the scale as dichotomous with the response "strongly agree" taken as a "yes" response and all other responses as a "no" response. In some cases the direction of the question was reversed and "strongly disagree" became the "yes" response.

Several combinations of items were tried on the board with the selection made on the basis of those questions having the best "cutting points" in terms of a wide range of popularity. In this way, items with increasing popularity may be thought of as making progressive cuts into or definitions of the attitude universe.

It was found, however, that two questions of similar popularity tended to interfere with each other and increase errors sharply. Therefore, items were selected to give a wide range in popularity. In most cases, too, only one item was selected to define each cutting point. The coefficients of reproducibility found for various combinations of items ranged from .82 to .88. The higher coefficients were gained only when the number of items was reduced drastically-to four or five items--and in these cases the scaled scores were skewed and heavily weighted to either the high or low ends of the distribution.

The difficulty encountered in scaling the sixteen items indicated that all sixteen items certainly do not define a single universe. Educational motivation is apparently a multi-dimensional concept and this judgement is confirmed in the findings of the study by Mitchell.

However, several of the combinations of items did qualify as quasi-scales in that their coefficients approached the arbitrary .90. Moreover, the range of item popularities and the nature of error patterns were

consistent with the criteria of quasi-scales. In this case, too, the scale score was not to be used to reproduce the response pattern of an individual in which case a high coefficient of reproducibility is essential. Rather the scale score was to be used as a variable for correlation with other external criteria. This use of quasi-scales is regarded as legitimate. (Stouffer, p. 162).

Rapid machine methods of scalogram analysis. The reproducibility coefficients reported above were, of course, derived from the sample of one hundred individuals from the total 2,069 respondents. Some concern was felt for the representativeness of this sample despite its random selection. Since the task of selecting other samples of one hundred and testing them in various combinations of items is extremely laborious and likely inconclusive, the summary method suggested by Green for scalogram analysis was adopted.

The advantage of the scalogram board, however, is that it permits a visual check upon the shape of the parallelogram, the patterns of errors, and the distribution of scale scores. Therefore, the sample of one hundred individuals was used to select the most promising items and these items were compared with their popularities in the total sample of 2,069. Several discrepancies were discovered indicating that the sample of one hundred was not completely representative of the total sample.

Eleven items representing the best items as indicated by the scalogram board and by item popularities in the total sample were selected to form the final scale.

The scale items. Since the data from the Pupil Questionnaires were

already punched on I. B. M. cards, Green's method of scalogram analysis was easily applied to the total sample of 2,069. The method involves the counting of errors through selective sorting of the items to be scaled. It then yields a coefficient of reproducibility for the items involved which has been shown to be very close to the true coefficient. The machine method gives a coefficient which is usually a slight underestimation of true reproducibility since only the most common error patterns are traced. Tracing of the rarer error patterns requires much more sorting for very little increased accuracy of estimation. Error patterns not traced are those involving three or more consecutive errors in an individual's score pattern. Analysis with the scalogram board revealed few such high order error patterns. In any case, the machine method would assign an error of two for each such pattern, thus reducing error in the estimated reproducibility to a minimum. (Green, p. 79).

The following eleven items are those chosen for analysis using the machine summary method. The figure in per cent following the item indicates its popularity in the sample of 2,069 pupils. That is, the figure indicates the percentage of those endorsing the item as "strongly agree" (SA) or, in the case of the items which were scored negatively, as "strongly disagree" (SD). The items are arranged here in ascending order of popularity. The item number refers to the numbering of the item on the original questionnaire.

(1) Item #7 High school is not enough education to be successful in life. (18 per cent SA).

(2) Item #11 Getting good marks in school is more important than being popular. (25 per cent SA).

(3) Item # 9 Working hard in school is more important than having a good time. (28 per cent SA).

(4) Item #16 Doing your homework is more important than going out to have a good time. (36 per cent SA).

(5) Item # 5 Getting a good education is worth more than making a lot of money. (36 per cent SA).

(6) Item # 4 Most teachers try to help students as much as they can. (42 per cent SA).

(7) Item # 6 Some students don't appreciate the opportunities they have in school. (48 per cent SA).

(8) Item # 8 School should not be compulsory after Grade 8. (49 per cent SD).

(9) Item # 3 A school is just as important in a community as a hospital. (54 per cent SA).

(10) Item # 1 All students should go to school for as many years as possible. (64 per cent SA).

(11) Item #2 Even though school is hard, it will help me to be a better person in later life. (76 per cent SA).

Although items #16 and #5 have the same popularity, evidence from the scale analysis using the scalogram board indicated that these two questions fitted reasonably well and served to improve the distribution of scale scores. Their moderate popularity may account for the fact that these two questions of similar popularity fitted well while questions of similar but extreme popularity fitted together poorly.

The reproducibility of these eleven questions calculated by Green's machin method was .83. This coefficient, together with the range of

popularity and the absence of high order errors as found on the board analysis indicates that these items may be considered to form a quasi-scale. The distribution of the scaled scores is given in Appendix B.

The Reporting of Findings

The findings of the study are reported in the next three chapters. The first of these chapters reports the findings in relation to the principal axes factoring of the image matrix P. The next chapter reports the findings in the analysis of variance of the scaled achievement scores, and the final chapter reports the findings of the analysis of variance of the achievement factor scores.

CHAPTER VII

FINDINGS RELATED TO IMAGE ANALYSIS AND PRINCIPAL AXES FACTORING

From the basic intercorrelation matrix, R, image analysis proceeds through the inversion of R to the derivation of the matrices S^2, Q, and P. Observed values for these matrices are given in this chapter as they were found in the analysis of the thirteen test variables used in the study. The image covariance matrix, P, may be factored using any of the conventional factor analytic techniques. In this study, P was factored using the principal axes method. Eigenvalues and eigenvectors derived from this factoring are also presented in this chapter as are their associated principal axes factors. Finally the results of the varimax rotation of the principal axes factors are presented. The seven rotated factors having loadings greater than .200 are interpreted.

Image analysis

In conventional factor analysis, tests are resolved into common and unique portions. In image analysis, test intercorrelations are also broken into common and unique portions--or as Guttman calls them, images and anti-images. This resolution of intercorrelations into common and unique portions is done prior to factoring in image analysis, whereas in factor analysis the definition of these portions may be accomplished only after factoring. Only the common portion of intercorrelations is factored in image analysis; in factor analysis the total intercorrelations are

factored. The required data for image analysis are provided by the observed correlation matrix and its inverse.

The intercorrelation matrix. The intercorrelations of the thirteen tests used in this study are presented in Table I. The three socio-economic tests have their highest correlations with each other and moderate to low correlations with the achievement and academic ability tests. The tests of study and motivation have their highest correlations with each other and low correlations with all other tests, as is reflected in their low communalities. The Home Index scale is the most sensitive of the three socio-economic tests to the achievement tests. The achievement tests and the academic ability tests have high correlations with each other. However, the achievement tests are apparently not separable from the ability tests on the basis of their intercorrelations with each other. The Social Studies and Science tests have the highest correlation observed in the matrix and seem to have much that is common to other tests. These tests have, in fact, the highest communalities in the battery. The Quantitative scale of SCAT seems to be the most independent test of the achievement and ability groups. It has the lowest communality in the battery with the exception of the socio-economic and motivation tests.

The inverse of the intercorrelation matrix. The inverse of the observed correlation matrix is presented in Table II. The elements of this matrix, as given in this table, contain only four figures although the matrix inversion program used yields elements to eleven figures. The re-inversion of this matrix gave a virtually exact reproduction of the original intercorrelation matrix. An exact inverse is essential for the

TABLE I

CORRELATION MATRIX OF THIRTEEN VARIABLES

(N = 2,069) Variable[a]	1	2	3	4	5	6	7	8	9	10	11	12	13	Mean	S.D.
1. S.E.S.	[b](252)	461	397	035	011	181	185	153	186	158	179	218	144	100.9	27.9
2. Home		(383)	507	034	033	307	295	241	289	238	267	349	220	11.02	3.65
3. Occup.			(308)	069	024	199	251	233	221	202	221	231	150	50.10	6.66
4. Study				(138)	237	-025	166	146	073	124	102	-032	026	3.21	1.06
5. Motiv.					(064)	-047	052	046	023	035	015	-028	-007	4.78	2.50
6. Rdg.						(696)	697	657	686	578	633	811	588	29.51	8.29
7. Lit.							(775)	831	780	677	712	718	611	28.92	8.15
8. Lang.								(740)	741	694	695	686	641	60.21	16.27
9. Soc. St.									(791)	741	842	718	667	58.86	16.22
10. Math.										(771)	813	580	781	60.62	16.54
11. Science											(793)	669	690	59.82	16.17
12. V score												(742)	617	34.42	10.82
13. Q score													(659)	34.54	8.59

[a]See text for full names of variables.

[b]Squared multiple correlation estimates of communality are in parentheses. Decimal points have been omitted in reporting intercorrelations.

TABLE II

THE INVERSE OF THE CORRELATION MATRIX OF THIRTEEN VARIABLES

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13
1. S.E.S.	1.336	-.4428	-.2908	-.0206	.0082	.0242	-.0005	.0632	-.0223	-.0032	-.0181	-.0907	-.0193
2. Home		1.621	-.5684	-.0098	-.0476	-.0994	-.0889	.1942	-.0724	-.0288	.0418	-.2876	.0000
3. Occup.			1.445	-.0257	.0111	.0609	-.0838	-.1597	.0588	-.0606	-.0736	.0181	.1164
4. Study				1.160	-.2324	.1333	-.3432	-.1356	.1366	-.1311	-.1151	.2490	.1324
5. Motiv.					1.068	.1143	-.0552	-.0359	-.0528	-.0490	.0617	.0091	.0359
6. Rdg.						3.290	-.5203	-.1227	-.2794	-.1136	.0261	-1.860	-.1262
7. Lit.							4.442	-2.036	-1.108	-.2264	.0045	-.5108	.1661
8. Lang.								3.845	-.3516	-.5318	.0615	-.4155	-.3089
9. Soc. St.									4.777	-.1728	-2.299	-.4312	-.1258
10. Math.										4.364	-1.980	.6404	-1.757
11. Science											4.834	-.4984	.0023
12. V score												3.878	-.5220
13. Q score													2.929

method of image analysis depends on information supplied from the elements of the inverse. Where possible, the fullest information from the inverse was used and rounding was kept to a minimum. In the discussion which follows, the elements of the inverse will be referred to as the s_{jk} where j is the jth row or test and k is the kth column or test.

From the s_{jj} (j = 1 to k), the squared multiple correlations of the n tests may be calculated. These diagonal elements of the inverse may be used to derive the squared multiple correlations as follows:

$$R_{j.k...n} = 1 - 1/s_{jj}.$$

The diagonal elements of the inverse are also used to form the matrix S^2. The reciprocals of the diagonal elements of the inverse, the the $1/s_{jj}$, form the diagonal elements of a new matrix S^2. Apart from these diagonal elements, all other elements in S^2 are zero. That is, S^2 is a diagonal matrix.

The Q matrix, the matrix of anti-images, is formed by the matrix multiplication of $S^2 R^{-1} S^2$. Each element in Q is thus formed by $s_{jk}/s_{jj}s_{kk}$ (j,k = 1 to n).

The inverse matrix may also be used to derive the matrix of weights for predicting a test from the remaining n - 1 tests in the battery. That is:

$$W = I - S^2 R^{-1},$$

where W is the matrix of weights and I is an identity matrix.

Very little that is of immediate consequence can be determined by the observation of the inverse matrix without further calculation. One important exception is pointed out by Guttman. The inverse of the matrix of observed correlations may be used as a direct test of the possible

existence of common factors.

> . . .It provides an entirely new way of testing empirical data for the possible existence of common factors. Given an observed matrix R_n, compute R_n^{-1} and see if the non-diagonal elements are all close to zero. Such a criterion requires no preliminary determination of "communalities" nor "fitting" of factor loadings, nor specification of the number of common factors. (Guttman, 1953, p. 282).

Unless these non-diagonal elements approach zero, any attempt to define the common factor space as less than n/2--the usual hypothesis of Thurstonian factor analysis--is impossible.

Examination of the non-diagonal elements of the observed inverse in Table II indicates that there are many non-zero elements. Thus, it may be expected that the number of factors required to describe the test space of the thirteen tests is n, the number of tests. This expectation is a consequence of the assumptions of image analysis.

The anti-image covariance matrix. Table III contains the matrix of anti-image covariances. The diagonal elements of this matrix are the unique portions of the thirteen tests. The non-diagonal elements are the "special pairwise linkages" between the errors of prediction in the thirteen tests. (See Guttman, 1953 and Chapter III of this report). In factor analysis, the assumption is that all these linkages are zero. It is clear from Table III that the thirteen tests observed in this study have many non-zero linkages. These linkages, as presented in the non-diagonal elements of Table III, may be either positive or negative. This situation arises from the partial correlation which may exist between the errors of prediction in tests. That is, when the correlation between the errors of prediction of two tests is held constant, a partial correlation exists between these errors which may be either positive or negative. (Guttman,

TABLE III

THE MATRIX OF ANTI-IMAGE COVARIANCES

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13
1. S.E.S.	[a]748	-204	-150	-013	005	005	000	012	-003	000	-002	-017	-004
2. Home		617	-242	-005	-027	-018	-012	031	-009	-004	005	-045	000
3. Occup.			692	-015	007	012	-013	-028	008	-009	-010	003	027
4. Study				862	-187	034	-066	-030	024	-025	-020	055	038
5. Motiv.					936	032	-011	-008	-010	-010	011	002	011
6. Rdg.						304	-035	-009	-017	-007	001	-145	-013
7. Lit.							225	-119	-052	-011	000	-029	012
8. Lang.								260	-019	-031	003	-027	-027
9. Soc. St.									209	-008	-099	-023	-008
10. Math.										229	-093	037	-137
11. Science											207	-026	000
12. V score												258	-045
13. Q score													341

[a]Decimal points omitted.

1953, p. 288).

As presented in Table III, the tests having the largest pairwise linkages are the three tests of socio-economic status and the tests of reading and study. Substantial linkages are also found among several of the tests of achievement and ability. All of these larger linkages are negative and will therefore act to reduce the commonness of the intercorrelation matrix. The intercorrelation matrix, therefore, implies a greater commonness between tests than is justified on the basis of anti-image covariances.

This anti-image matrix is the matrix Q in Kaiser's notation. Adding R and Q and subtracting the matrix $2S^2$ produces P, the image covariance matrix.

The image covariance matrix. The image covariance matrix, P, is presented in Table IV. This matrix contains the squared multiple correlation estimates of communalities in its diagonal and the modified test intercorrelations in its off-diagonals. The matrix is Gramian and therefore the image covariances reproduce the communalities exactly and any factors extracted will be "real", non-negative entities.

The image covariances show patterns of relationships which are similar to those observed in the test intercorrelations presented in Table I. However, the educational ability tests stand out with greater clarity from the achievement tests although there is still a close relationship indicated between them. The socio-economic, study, and motivation variables still show the same relative independence which was apparent in the test intercorrelations. This P matrix of image covariances

TABLE IV

THE MATRIX OF IMAGE COVARIANCES

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13
1. S.E.S.	[a]252	257	247	022	016	186	185	165	183	158	177	201	140
2. Home		383	265	029	006	289	283	272	280	234	272	304	220
3. Occup.			308	054	031	211	238	205	229	193	211	234	177
4. Study				138	050	009	100	116	097	099	082	023	064
5. Motiv.					064	-015	041	038	013	025	026	-026	004
6. Rdg.						696	662	648	669	571	634	666	575
7. Lit.							775	712	728	666	712	689	623
8. Lang.								740	722	663	698	659	614
9. Soc. St.									791	733	743	695	659
10. Math.										771	720	617	644
11. Science											793	643	690
12. V score												742	572
13. Q score													659

[a]Decimal points omitted.

was factored using the principal axes method.

Principal Axes Factoring

Principal axes factoring yields a series of eigenvalues and eigenvectors for the equation, $Px = \lambda x$, where P is the image covariance matrix, x are the eigenvectors, and λ are the eigenvalues. The eigenvectors are normalized vectors--the sum of the squares of their loadings is unity and the product of any two vectors is zero. The λ's are indications of the size of the associated principal axes factors. Re-scaling of the eigenvectors by the square root of its associated eigenvalue, produces a principal axis factor.

Eigenvalues and eigenvectors. The eigenvalues and eigenvectors derived from the principal axes factoring of the image covariance matrix, P are presented in Table V. The sum of the test communalities estimated in the P matrix is 7.112. The sum of the eigenvalues derived in the factoring of P is 7.112. The convergence of these two figures is an indication of the accuracy of the estimate of communality. In conventional factor analysis, the estimate of communality and the sum of the eigenvalues can be made to "fit" only by varying the number of factors extracted until convergence is reached. Even in this case, there is no proof that the estimates will converge and varying the number of factors involves problems of estimating the significance of factors. Moreover, in conventional factor analysis, the estimating of the number of factors means that the nature of the factors extracted depends on the number of factors extracted. Rotation of these factors involves further indeterminacies, since the nature of the rotated factors also depends on the

TABLE V

EIGENVECTORS AND EIGENVALUES FROM THE IMAGE COVARIANCE MATRIX

Variable	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	XIII
1. S.E.S.	1018	5229	0865	-1464	0529	-1801	0067	2597	4046	-0015	6308	1385	0598
2. Home	1518	5772	-0186	-1030	-1686	6865	1751	-1548	-2436	-0508	-1084	0646	-0208
3. Occup.	1220	5420	2026	0006	0946	-5151	-2220	-0640	-1399	-0021	-4925	-2262	0951
4. Study	0396	-0024	5318	4282	0373	1549	-3779	-3071	-0063	0078	2682	-1410	-4143
5. Motiv.	0082	0221	3089	2565	-1887	-1387	3698	3772	0066	-1377	-2646	5095	-3943
6. Rdg.	3270	0287	4709	1057	-2656	-0667	-2690	4220	-2898	-0121	1262	-2581	-4086
7. Lit.	3557	-0332	-0033	4007	-1137	-2602	5007	-2612	-3353	-0439	3089	-0987	3040
8. Lang.	3477	-0871	0527	3962	-1355	2605	-2038	2474	4519	-2643	-2218	-1044	4340
9. Soc. St.	3659	-0902	0578	0058	2701	0684	-2794	1512	-2811	5364	-0046	5015	2399
10. Math.	3419	-1807	3273	-3283	5349	1267	1936	2680	-1588	-3199	0272	-2900	-1065
11. Science	3585	-1289	1818	-3013	-3553	-0188	2520	-0772	3431	5500	-1375	-2680	-1543
12. V score	3383	0602	-4402	1385	4401	-0484	0957	-3885	3622	-0667	-1456	1914	-3451
13. Q score	3189	-1684	1040	-4147	3764	-1551	-2956	-3293	-0576	-4508	0650	3383	0013
Eigenvalue	5.667	.6115	.2860	.1698	.0984	.0823	.0539	.0472	.0304	.0274	.0223	.0111	.0042

number of factors rotated.

Since factor analysis deals with variables in standardized form, the total variance of the tests in the intercorrelation matrix is 13.000. Since the sum of the communalities estimated and the sum of the eigenvalues are both 7.112, the ratio of common variance to total variance is .547. Thus common variance is 54.7 per cent of total variance. A large portion of total variance in the test battery is thus unique variance; the most useful measure of the size of a factor, therefore, is the percentage of common variance which it accounts for.

As indicated in Table V, the number of common factors extracted is thirteen. These vary in size from one very large factor to several very small ones. These vectors resemble the results which might be obtained under the Spearman hypothesis of factor analysis where the communality of each test is estimated as unity. Such a matrix is Gramian and produces the same number of positive eigenvectors and eigenvalues as the number of tests. The image matrix P, however, has better estimates of communality and still preserves the Gramian properties of the matrix. Moreover, the nature of the factors extracted in the two methods would be different.

The factor matrix. The matrix of common factor loadings is presented in Table VI. The first common factor is a large general factor running through virtually all tests. All loadings on this factor are positive with the highest loadings appearing for the achievement tests. Lower, though still high, loadings appear for the tests of academic ability. The tests of socio-economic status have moderate to low loadings

TABLE VI

FACTOR MATRIX OF THIRTEEN PRINCIPAL AXES FACTORS

Variable	Common Factor Loadings													
	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	XIII	H^2
1. S.E.S.	2423	4089	0463	-0603	0166	-0517	0015	0564	0705	-0002	0943	0146	--38	252
2. Home	3615	4514	-0099	-0424	-0529	1970	0406	-0336	-0424	-0084	-0162	0068	-0013	383
3. Occup.	2904	4238	1083	0002	0296	-1478	-0515	-0139	-0244	-0003	-0736	-0239	0061	308
4. Study	0944	-0019	2844	1765	0017	0444	-0877	-0667	-0011	0129	0401	-0148	-0268	138
5. Motiv.	0196	0173	1652	1057	-0592	-0398	0859	0820	0011	-0228	-0395	0538	-0255	064
6. Rdg.	7785	0224	-2518	0436	-0833	-0191	-0625	0917	-0505	-0020	0188	-0272	-0265	696
7. Lit.	8469	-0259	-0018	1651	-0357	-0746	1163	-0568	-0584	-0072	0462	-0104	0197	775
8. Lang.	8277	0681	0281	1633	-0425	0747	-0473	0538	0787	-0437	-0331	-0110	0281	740
9. Soc.St.	8711	-0705	0309	0023	0847	0196	-0649	0328	-0490	0888	-0006	0529	0155	791
10. Math.	8139	-1413	1751	-1353	1678	0363	0449	0582	-0276	-0529	0040	-0306	-0069	771
11. Science	8535	-1008	0972	-1241	-1114	0053	0585	-0167	0598	0910	-0205	-0283	-0100	793
12. V score	8054	0470	-2354	0571	1380	-0139	0222	-0844	0631	-0110	-0217	0202	-0223	742
13. Q score	7592	-1317	0556	-1709	-1181	-0445	-0686	-0716	-0100	-0746	0097	0357	0000	659
Sum of squares	5.677	.6115	.2860	.1698	.0984	.0823	.0539	.0472	.0304	.0274	.0223	.0111	.0042	7.112
% of common variance	79.68	8.60	4.02	2.37	1.38	1.16	.76	.66	.43	.39	.31	.16	.06	
% of total variance	43.59	4.70	2.20	1.31	.76	.63	.41	.36	.23	.21	.17	.09	.03	

Decimal points omitted in factor loadings and communalities (h^2).

on this factor and the tests of study and motivation have very low loadings.

The second factor is apparently a socio-economic factor since the three socio-economic tests have the highest loadings on it and all other test loadings are low. The next two factors have their highest loadings on the study and motivation tests and on the verbal and quantitative scales of SCAT. The principal axes factoring thus produces factors similar to those hypothesized for this study.

The Rotation of Axes

To assist in the interpretation of these factors, the principal axes position of the factors was rotated to approach the parsimonious representation found in simple structure. The varimax criterion was used for this rotation.

The varimax rotation. The result of the varimax rotation is presented in Table VII. Several rotations of varying numbers of the principal axes factors were performed. In the first rotation, all thirteen factors were used. Subsequent rotations used successively fewer factors until only the four largest factors were used. The rotation of the four largest factors only produced substantially the same results as the rotation of all thirteen factors. That is, the elimination of factors from the matrix of factors to be rotated did not affect the nature of the resulting rotated factors. Thus another advantage of image analysis is apparent: the results of rotations are seemingly invariant over the number of factors rotated. This condition does not hold in conventional factor analysis where the results of rotation are dependent upon the number of

TABLE VII

SEVEN ROTATED FACTORS HAVING LOADINGS GREATER THAN .200

Variable	Common Factor Loadings I	II	III	IV	V	VI	VII	h^2
1. S.E.S.		479						252
2. Home	205	500				298		383
3. Occup.		509						308
4. Study			350					138
5. Motiv.							243	064
6. Rdg.	754	234						696
7. Lit.	809	218						775
8. Lang.	800							740
9. Soc.St.	842	201						791
10. Math.	800				303			771
11. Science	837							793
12. V score	764	270		-219				741
13. Q score	759			233				659
Sum of squares	5.146	1.056	.217	.130	.136	.105	.120	
% of common variance	72.35	14.85	3.05	1.83	1.91	1.48	1.69	
% of total variance	39.58	8.12	1.67	1.00	1.05	.81	.92	

Decimal points omitted in common factor loadings and communalities (h^2).

factors rotated.

The factors presented in Table VII are the seven largest derived from the rotation of all thirteen factors. Loadings below .200 have been omitted to facilitate the interpretation of the factors. The communalities observed for the rotated factors are the same as those observed for the principal axes factors and for the estimates of communalities in the image covariance matrix. This invariance of test communalities over rotation is a characteristic of the varimax rotation and other orthogonal analytic rotation methods.

The size of the factors is, however, modified somewhat in rotation as is apparent from the comparison of the sums of squares of the rotated factors with the sums of the squares of the principal axes factors. The large first factor is reduced in size and the second factor increased in size. Since communalities are invariant over rotation, whatever variance is taken from one factor must be added to others.

Interpretation of Factors

The rotated factors seem to represent the same qualities appearing in the principal axes factors but their structures are simpler.

Factor I has its highest loadings on the five Alberta Departmental Examinations. The three tests of educational ability retain their high loadings on this factor too, but are lower than those obtained for the tests of subject matter knowledge. The only other test having a loading on this factor is the Home Index scale which has a moderate loading. This factor would have been more satisfactory for the purposes of this

study if the ability tests (SCAT: Verbal; SCAT: Quantitative; STEP: Reading) had lower loadings than those observed. However the nature of these tests apparently precludes this. The SCAT is designed as a predictor of educational achievement, though it is not as specific as standardized achievement tests. However, it apparently has more in common with such achievement tests than it does with tests such as the Progressive Matrices (Elley; Hukins) which is relatively independent of achievement and is a purer measure of general ability and Spearman's g. The loadings of this factor are, however, consistent with those of a v:ed factor and this factor may be identified as achievement subject to development through school learning. This factor alone accounts for 72.3 per cent of common variance.

The Elley and Hukins studies both involved a factor analysis of tests of achievement and intellectual ability. In Elley the test battery included seven "culture-reduced" tests of intellectual ability and several standardized achievement tests of vocabulary, reading comprehension, arithmetic, spelling, language, and verbal ability. The largest factor extracted from this battery was a factor identified as Spearman's g. The second largest factor, identified as a v:ed factor, had highest loadings on the tests of achievement and loadings below .200 on each of the seven culture-reduced intelligence tests. (Elley, p. 117). The v:ed factor identified by Hukins contains moderate to high loadings on tests of science knowledge. The two SCAT sub-tests both load on this factor with the SCAT: Verbal having the highest loading. The Progressive Matrices test of intellectual ability had a loading of .05. (Hukins, p. 109).

The second factor has highest loadings on the three tests of socio-economic status. It accounts for 14.8 per cent of common variance and may be identified as a socio-economic factor. It is apparently sensitive to verbal tests since the STEP: Reading, SCAT: Verbal, Literature, and Social Studies tests have moderate loadings on it. It is not, however, a measure of verbal ability as such since the occupation scale, which has the highest loading on the factor, is not a performance test but a ranking assigned to an individual on the basis of status achieved. The verbal aspect of the factor may therefore be thought of as a measure of a literate quality associated with socio-economic status.

The remaining factors are very small and interpretation of them is difficult and perhaps uncertain. Three of them will, however, be named and some justification given for this in the following section of this report.

The Factor III has a single loading on the test of study hours. It is a measure of the self-reported industry of a student. In this rotation it is apparently quite independent of achievement and socio-economic status although in its principal axis position it was associated with the Motivation test and with a negative score on the SCAT: Verbal.

Factor IV has two small loadings on the SCAT: Verbal and the SCAT: Quantitative tests. These loadings are opposite in direction and similar in size. It may be regarded as the beginning of a general ability or g factor. However, the battery of tests used here apparently does not contain enough tests to establish this factor strongly.

Factors V and VI have single loadings on the Mathematics and Home Index tests respectively. These factors will not be named.

Factor VII has a single loading on the Motivation scale and may be interpreted as an aspiration level of an individual with or without attendant scholastic achievement or industry.

Discussion

The strength of the factors identified beyond Factor I and Factor II presents a difficulty in relating the findings of the factor analysis to the hypotheses of the study. One justification for interpreting these factors may be found in the fact that these factors are the result of image analysis rather than conventional factor analytic techniques. The factors derived from conventional factor analysis are hypothetical random variables whereas in image analysis they are concrete observable entities. Concern for the size of factors in conventional factor analysis is thus partly the result of caution stemming from uncertainty about the precision of the factors which have been extracted. Despite the concreteness of the small image factors, it was decided that the size of these factors did not warrant the acceptance of the hypotheses predicting the motivation factor and academic ability factor. The beginnings of such factors are, however, indicated and it may be presumed that other studies using clearer tests of these qualities would identify stronger motivation and ability factors.

The first two factors, however, are clearly identified even by the standards of conventional factor analysis. The use of these two factors is therefore justified in the analysis to be reported in subsequent parts of this study. The first factor, the v:ed factor, is independent of the important socio-economic factor. To fit the design of the study it ought

also to be independent of an academic ability factor. Since there is no strong academic ability factor, this independence is not certain. However, the factor does resemble other v:ed factors which have been extracted from batteries of achievement and ability tests which included tests with high g loadings. (Elley: Hukins). The first factor may, however, contain some variance due to academic ability.

One of the strengths of factor analysis is its potential for partitioning variance into common and unique portions. The first factor may be regarded therefore as a more precise measure of achievement than the total achievement scores. Moreover, the resolution of tests into common variance distributes this variance among common factors. More of the variance from the ability tests has been distributed to unique factors and to factors other than Factor I than has variance from the achievement tests. This distribution is seen in the communalities and in high loadings of the achievement tests on Factor I.

Moreover, the interpretation of Factor I depends upon the relationship of achievement to ability. There is a considerable overlap between achievement and ability as indicated in the correlations between the five achievement tests and the three ability tests. Cronbach indicates that the Kuhlmann-Anderson ability test has 71 per cent overlapping variance with achievement. Unreliability and uniqueness accounts for 14 per cent of its variance and only 15 per cent may be regarded as ability variance distinct from achievement. (Cronbach, p. 224). The SCAT tests have a similar high overlap with achievement. The question of interpretation arises then in deciding whether the overlapping variance is due to the

achievement or the ability of an individual. Since the 15 per cent of variance is comparable to the innate, ineducable g, the 71 per cent of variance is associated with a learned and educable achievement.

These considerations, then, strengthen the interpretation of Factor I as a factor showing relative independence from academic ability and considerable independence of socio-economic status.

Test selection. The definition of the factors identified in this study can be improved by altering the selection of tests used to define the factors. Both the Study and Motivation tests used in this study have large unique factors. This uniqueness is due both to unreliability and to specific factors in the tests. The amount of each of these sources of uniqueness is uncertain. One study of educational motivation does throw some light on the nature of motivation. (Mitchell). This study concludes that educational motivation is not a unitary concept but a multidimensional entity. Only one of six factors in educational motivation which are identified in the study has any relation to academic achievement. This factor is identified as Academic Motivation and Efficiency and is made up of a variable representing the discrepancy between achievement and potentiality and two variables reflecting the actual effort expended to achieve an effective academic performance. (Mitchell, p. 184). It may be presumed therefore that the indices of motivation used in this study represent specific factors unrelated to ability or achievement.

Guttman indicates how tests may be selected so as to improve the definition of factors or to increase the communality of a test. The

method depends upon knowledge of the multiple correlation of a test with other tests.

If tests are added to a bettery of tests with rank A and if these tests have substantial correlation with the tests already in the battery, the rank of the matrix (the number of factors in it) and the communalities of it are unaffected.

> . . .In order to keep these additional tests hitting on almost all r cylinders, so to speak, where r is the number of factors, it is in general necessary for them to have sustantial correlations with all of the original tests. This helps explain the known fact that adding to a regression tests that have substantial correlations with the tests already in the regression will not raise a multiple correlation coefficient much. (Guttman, 1940, p. 98).

On the other hand, if independent tests are added to the matrix, a better definition of factors is obtained and communalities are increased. New factors may also be established.

> If, however, tests are added which correlate substantially with the "dependent" test but very slightly with the other tests in the regression, these new tests tend to bring new common factors in, thus increasing the rank of A and increasing the communality of the "dependent" test. This raises the upper bound for the multiple correlation coefficient and enables it to increase more readily. Therefore, if it is desired to obtain better estimates of a given test, those tests which correlate substantially with that test but correlate negligibly amongst themselves are the most desirable as "independent" tests. (Guttman, 1940, p. 99).

Better definition of the academic ability factor could be obtained, therefore, by including in the test battery purer measures of Spearman's g such as the Progressive Matrices Test and other "culture-reduced" tests of intellectual ability.

The Hypotheses

Four of the research hypotheses of the study predicted factors

which were to be found in the matrix of the thirteen test intercorrelations.

Hypothesis 4. The intercorrelation matrix can be factored to yield a factor on which the Alberta Departmental Examinations of subject knowledge have the highest loadings.

The largest factor extracted from the matrix, identified as a v:ed factor was consistent with this hypothesis. The three ability tests in the matrix also had large loadings on this factor, however. After consideration of the meaning of these ability test loadings, the hypothesis was accepted. The ability tests must therefore be regarded as measures of achievement primarily rather than as measures of general intellectual ability.

This factor accounted for the major portion of common variance in the matrix. Its identification as an achievement factor permitted the derivation of individual factor scores. These achievement factor scores were then available for use in subsequent hypotheses of the study.

Hypothesis 5. The intercorrelation matrix can be factored to yield a factor on which tests of socio-economic status have the highest loadings.

The second factor extracted from the matrix was identified as a socio-economic factor consistent with this hypothesis. Each of the three socio-economic tests had their highest loadings on this factor. The hypothesis was therefore accepted.

Since the axes of principal component factors are orthogonal, the factor scores of Factor I, the achievement factor, are independent of

factor scores on the socio-economic factor. In this way, the achievement factor scores are free of the effects of socio-economic status upon achievement.

Factor II, the socio-economic factor, showed sensitivity to certain verbal tests of achievement and ability. It may thus be thought of as representing the socio-economic components of these tests. The factor was considerably smaller than the achievement factor but was clearly defined and of significant size.

Hypothesis 6. The intercorrelation matrix can be factored to yield a factor on which tests of educational motivation have the largest loadings.

The test of study and the test of Motivation finally loaded not on one factor but on two with each test defining a separate factor. Both of these factors were very small and, although identified, they were not regarded as clear enough or large enough to accept this hypothesis. Both of these tests of motivation had large unique loadings which may be attributed to their unreliability and to specific factors which are independent of other tests and factors in the battery.

Hypothesis 7. The intercorrelation matrix can be factored to yield a factor on which tests of verbal and quantitative ability have the highest loadings.

The SCAT: Quantitative and the SCAT: Verbal scales and the STEP: Reading test failed to define a single factor representing general intellectual ability. Their high correlations with the achievement tests apparently precluded the establishment of such a factor. However, the Quantitative

and Verbal scales did load equally and with opposite signs on a small factor. This factor may be regarded as the beginning of a general intellectual factor. This factor was too small and ill-defined to warrant accepting this hypothesis.

Summary

The most important factor defined in the factoring of the matrix of test intercorrelations was the factor identified as general achievement. This factor was shown to be independent of a second factor representing the socio-economic component of the test battery. Two other factors representing motivation and general intellectual ability did not appear in the factoring. On the basis of this study, therefore, the independence of the achievement factor from motivation and intellectual ability is not established. However, the achievement factor is similar in size and configuration to similar achievement factors identified in other studies. These studies identified as well a general intellectual factor. The tests having large loadings on this factor were culture-reduced tests yielding the purest measures of Spearman's g. Significantly these same tests had virutally zero loadings on the achievement factor.

The identification of the achievement factor in this study is important also for further analysis and testing of hypotheses. The individual factor scores of this factor are measures of the amount of the factor possessed by an individual. In this form, these scores are free of certain of the errors found in the total achievement score since the unique and specific factors in a test are not part of the score.

Moreover, these scores are free of the socio-economic factor as indicated by the independence of the achievement factor and the socio-economic factor. It may be presumed that these scores are also independent of the general intellectual factor as indicated by the independence of achievement and intellectual factors derived in other studies.

CHAPTER VIII

FINDINGS RELATED TO THE ANALYSIS OF VARIANCE OF ACHIEVEMENT SCORES

The achievement scores of the 2,069 pupils in the sample were subjected to analysis of variance as outlined in Chapter VI. The scores analyzed by this technique are the results obtained by the 2,069 pupils on the Alberta Departmental Examinations. These examinations and the scores obtained from them are described in Chapter V on Test Data; these scores are denoted in this study as "raw achievement scores" to differentiate them from the factor achievement scores which are derived from them. In fact, however, these scores are not the raw scores obtained by the pupils on the examinations but the scaled scores assigned by the Alberta Department of Education in accordance with their policy of "examinee standards." (MacArthur and Hunka, pp. 6ff).

For the purposes of this study, these Departmental Examination scores are basic data; they are thought of as "raw" only in relation to the factor scores derived from them. Both these scaled scores and the factor scores have an internal dependence imposed by the scaling and by the standardizing of factor scores. As a result, both sets of achievement scores violate the assumption of independence in analysis of variance. The consequences of the violation of this and possibly other assumptions in the analysis of variance are also discussed in the section on analysis of variance in Chapter VI.

Achievement scores for pupils in the sample were obtained on each of five Departmental Examinations of subject matter knowledge. These examinations are English Literature, English Language, Social Studies,

Mathematics, and Science. This chapter reports the results of the analysis of variance performed on the scaled scores obtained from each of these examinations. In this way, the total variation in each examination may be resolved into sources from the class, school, and district in which pupils are located. The residual source is considered as pupil variation within classes. This analysis permits the testing of the first four hypotheses of the study and permits variation from the sources described above to be expressed as percentages of the total variation.

The Analysis of Variance

The analysis of variance is based upon the model given in Appendix A. The form of the analysis of variance based on this model is given in Table VIII. As indicated in this table, total variation has sources in classes, schools, districts, and residuals. The sums of squares are designated as S_c for classes, S_s for schools, S_d for districts, S_r for residuals, and S_t for the total. Degrees of freedom for each source of variation depend upon the total number of pupils in the sample, the number of districts, the number of schools in a district, and the number of classes in a school. The formation of mean squares and F-ratios are also indicated in the table. These ratios involve the comparison of within and between variance. The residual is the single source of within variance while between variance has sources in classes, schools, and districts.

The achievement scores from each of the five Departmental Examinations were analyzed according to the analysis of variance given in Appendix A and the results expressed according to the form given in Table VIII.

TABLE VIII

THE ANALYSIS OF VARIANCE

Source of variation	SS[a]	d of f[b]	MS	F
Between classes	S_c	$IJ(K - 1)$	$S_c/IJ(K - 1)$	MS_c/MS_r
Between schools	S_s	$I(J - 1)$	$S_s/I(J- 1)$	MS_s/MS_r
Between districts	S_d	$I - 1$	$S_d/(I - 1)$	MS_d/MS_r
Residual	S_r	$N - IJK$	$S_r/(N - IJK)$	--
Total	S_t	$N - 1$	--	--

[a]The subscripts in this and other columns refer to classes, schools, districts, and residuals.

[b]I = the number of districts,

J = the number of schools in a district,

K = the number of classes in a school,

N = the total number of pupils in the sample.

The same notation is used in the Mean Square column.

Sources of variation in achievement. Tables IX to XIII summarize the results of the analyses of variance performed on the five Grade IX Departmental Examinations in English Literature, English Language, Social Studies, Mathematics, and Science. Each of these analyses involved the same sample of 2,069 pupils in the 22 districts, 44 schools, and 88 classrooms participating in the study. Three F - tests were performed for each subject. The mean square for residuals was used in the denominator of each of these tests; the numerators were formed from the mean squares for classes, schools, and districts. The mean square for residual is made up of variance within classes while the other mean squares are made up of variances between classes, between schools, and between districts. F - tests formed in this way compare variance within classes with variance between classes.

Between class variance arises from differences between mean class achievement and the mean of all achievement in a school. If the mean achievements of classes in a school are equal, the only source of between class variance is error. If the equality of mean class achievements holds in all schools, F - ratios comparing between class variance and within class variance are non-significant.

Between school variance arises from differences between mean school achievement and the mean of all achievement in a district. If school means vary little from their district means, the only source of between school variance is error. Similarly, district variance arises from differences between mean district achievement and the grand mean of all achievement in the sample. If district means vary little from the grand mean, error is again the only source of between district variance, and

TABLE IX

THE ANALYSIS OF VARIANCE OF ACHIEVEMENT IN ENGLISH LITERATURE

Source of variation	SS	Per cent of variation	d of f	MS	F
Between classes	23,687.8	17.22%	44	538.4	12.3[a]
Between schools	8,686.3	6.31%	22	394.8	9.0[a]
Between districts	18,296.5	13.30%	21	871.3	19.8[a]
Residual	86,923.5	63.17%	1981	43.9	--
Total	137,594.1	100.00%	2068	--	--

[a]Significant above .01.

TABLE X

THE ANALYSIS OF VARIANCE OF ACHIEVEMENT IN ENGLISH LANGUAGE

Source of variation	SS	Per cent of variation	d of f	MS	F
Between classes	92,596.0	16.90%	44	2,104.4	11.6[a]
Between schools	26,386.7	4.82%	22	1,199.4	6.6[a]
Between districts	68,804.6	12.56%	21	3,276.4	18.0[a]
Residual	360,011.4	67.72%	1,981	181.7	--
Total	547,798.7	100.00%	2,068	--	--

[a]Significant above .01.

TABLE XI

THE ANALYSIS OF VARIANCE OF ACHIEVEMENT IN SOCIAL STUDIES

Source of variation	SS	Per cent of variation	d of f	MS	F
Between classes	100,000.0	18.38%	44	2,272.7	12.5[a]
Between schools	26,072.4	4.79%	22	1,185.1	6.5[a]
Between districts	57,120.2	10.50%	21	2,720.0	14.9[a]
Residual	360.961.5	66.33%	1,981	182.2	--
Total	544,154.1	100.00%	2,068	--	--

[a]Significant above .01.

TABLE XII

ANALYSIS OF VARIANCE OF ACHIEVEMENT IN MATHEMATICS

Source of variation	SS	Per cent of variation	d of f	MS	F
Between classes	80,854.5	14.29%	44	1,837.6	10.2[a]
Between schools	39,076.5	6.90%	22	1,776.2	9.8[a]
Between districts	87,946.8	15.54%	21	4,187.9	23.2[a]
Residual	358,073.4	63.27%	1,981	180.7	--
Total	565,951.2	100.00%	2,068	--	--

[a]Significant above .01.

TABLE XIII

THE ANALYSIS OF VARIANCE OF ACHIEVEMENT IN SCIENCE

Source of variation	SS	Per cent of variation	d of f	MS	F
Between classes	97,993.6	18.11%	44	2,227.1	11.7[a]
Between schools	21,904.2	4.05%	22	995.6	5.2[a]
Between districts	45,760.4	8.46%	21	2,179.1	11.5[a]
Residual	375.374.6	69.38%	1,981	189.5	--
Total	541,032.9	100.00%	2,068	--	--

[a]Significant above .01.

F - ratios are again non-significant.

All the F - ratios reported in Tables VIII to XIII are significant above the .01 level. The three sources of between variance--classes, schools, and districts--are thus significantly larger than variance within classes.

Discussion

The largest source of variation in all five analyses of pupil achievement is the residual--the variation of pupils within classes. In these five tests of achievement, residual variation accounts for sixty-three to sixty-nine per cent of total variation. The F - tests depend, of course, not only upon this variation (sums of squares) but upon the degrees of freedom associated with it and other sources of variation. However, it may be useful to observe the various sources of variation as percentages of total variation since these sources have the important properties of independence and additivity. The independence arises from the fact that:

> if many samples of the same size from the same population are examined, the value of one of these components in a sample is in no way predictive of the value of the other.
> . . .Unless the components of variance are independent, the ratio of the related variances does not have an F distribution. (Walker and Lev, p. 210).

Thus knowledge of a within component is not predictive of a between component.

However, the additivity property of sources of variation is also a basic relationship in analysis of variance. This property permits

> the total variation described by the sum of squares of all N cases around the total mean [to be] broken up into components each

> of which is expressive of a type of variation. Thus it is the sum of squares rather than the variance which is analyzed. (Walker and Lev, p. 211).

Thus knowledge of the total sum of squares and the within sum of squares determines the between sum of squares.

In the analysis of variance model used in this study, the residual variation is the sum of squares of pupils within classes; this source is also the denominator of all the F - ratios formed. If residual variation is large in relation to class, school, and district variation, F - ratios will tend to be non-significant. Conversely, residual variation must be small in relation to between variation for F - ratios to be significant. This study predicts significant F - ratios formed with the three between sources of variation in classes, schools, and districts and the residual within variation from pupils. That is, the study predicts that educational systems have characteristics which tend to increase between variations.

It is thus reasonable to ask what contributes to the different sources of achievement variation. The residual variation of pupils in classes may be seen as being composed of a variety of pupil differences such as intelligence,previous learning, socio-economic status, motivation, and attitude. Errors of measurement and specific factors in achievement tests also contribute to such variation. Characteristics of classes contribute to between class variation. These characteristics might include the competence and training of teachers, the methodology of instruction in a subject, grouping of pupils, size of class, and the facilities available in a classroom for instruction. Between school variation is associated with such characteristics as school organization

and administration, library resources, the age and adequacy of the school building, and the characteristics and competence of principals. Between district variation would be associated with the economic resources of the district, expenditures on education, and community attitudes and values. The effectiveness of educational systems is thus dependent upon there being variation between classes, schools, and districts over and above that which is due to the within variation of pupils in classes.

System variations also depend upon differences in the characteristics of school, classes, and districts. The size of sample in this study and the variety of characteristics of the classes, schools, and districts indicates that such differences are present. The significance of the F - ratios indicates that these differences in school systems have effects which are greater than those which would be expected from the variability of pupils within classes.

The importance of the significant F - ratios may be indicated if it is asked, "What conditions in total achievement variation would produce non-significant F - ratios?" As already indicated, when pupil variation within classes is large in relation to the between variations, F - ratios are non-significant. This condition would indicate that total achievement variations are largely a function of individual differences in pupils. These differences, like those in the intelligence, socio-economic status, and attitudes of pupils, are, for the most part, beyond the control of educational systems. If these differences are the only source of achievement variation, the characteristics of educational systems can have little effect upon their major output--knowledge gained by pupils. The significance of the F - ratios reported in this study

indicates the presence of system effects upon the achievement of pupils. Or more precisely, the significant ratios indicate that each of the sub-systems in the educational system considered has significant effects upon achievement. The sub-systems are the classes, schools, and districts and the arrangements and characteristics of them.

The contribution of each of these sub-systems to total variation is indicated if the conditions that would establish the null hypotheses are considered. The null hypotheses of the analysis of variance state that there are no differences between class means and the school means, between school means and district means, and between district means and the grand mean of the sample. Thus for the null hypotheses to be established, means of all classes, schools, districts, and the grand mean itself must be virtually identical. The largest source of between variance would be error and the F - ratios would be non-significant. Whether a pupil learned under the conditions of one class, school, and district or under those of others would have no influence upon his achievement.

The indications of this study are that up to thirty-seven per cent of total pupil variation in examinations of subject matter knowledge is associated with the effects of classes, schools, and districts. This, then, is approximately the proportion of achievement variation which is subject to control by organizational arrangements in school systems. The major proportion of variation is, of course, due to the individual differences of pupils within classes. Moreover, the thirty-one to thirty-seven per cent of between variation is shared among the three sub-systems. There are also differences between subjects in the amount of variation associated with system effects and pupil effects.

Table XIV summarizes the proportions of variation associated with each source in the five academic subjects tested by the Alberta Grade IX Departmental Examinations. The distribution of proportions of total variation among the various sources is similar in all five subjects. Science has the largest proportion of total variation assigned to within variation and English Literature the least. This distribution may suggest that Science is the least susceptible to system effects and English Literature the most susceptible. In general, class differences account for the largest part of total between variation, district differences for the next largest part, and school differences for the least.

Although the three between variances are significantly greater than the within variances in all five subjects, in terms of variation the within component is large in comparison with the between component. This situation may explain some of the problems in educational research. The major portion of variation is due to pupil differences and therefore studies which attempt to relate economic inputs, teacher characteristics, teaching methods, and other system variables to achievement must also account for a large source of variation in pupils which would otherwise mask relationships. If the contributions of classes, schools, and districts to between variation alone is considered, the proportions of variation associated with each between source is much larger. The need to "partial out" pupil variation is thus clear.

The large proportion of between variation associated with classes as compared with schools and districts suggests too, that the organizational variables and characteristics of classes are those which have the greatest influence upon achievement. The importance of variables such as

TABLE XIV

PROPORTIONS OF BETWEEN AND WITHIN VARIATION ASSOCIATED WITH SOURCES IN FIVE ACADEMIC SUBJECTS

Source of variation	Subjects				
	English Literature	English Language	Social Studies	Mathe-matics	Science
Residual (within)	63.17%	65.72%	66.33%	63.27%	69.38%
Total between	36.83%	34.28%	33.67%	36.73%	30.62%
Between classes	17.22%	16.90%	18.38%	14.29%	18.11%
Between schools	6.31%	4.82%	4.79%	6.90%	4.05%
Between districts	13.30%	12.56%	10.50%	15.54%	8.46%

teacher competence, teaching methods, classroom organization, and group attitudes are thus indicated. The influence of any one of these must be smaller than class variation as a whole and the problems of isolating these effects will therefore be more difficult than measuring total effects on classroom variation.

It is perhaps not surprising that the classroom sub-system as the one closest to the actual scene of learning is the sub-system having greatest impact on achievement. It is perhaps significant then that the most remote sub-system, the district, also has a considerable effect. The importance of economic support for education in districts is thus indicated.

Hypotheses

Three hypotheses of the study are related to the analysis of variance of the raw achievement scores from the five tests of subject matter knowledge. Scores from each of the Grade IX Departmental Examinations were subjected to the analysis of variance. The analysis of variance hypotheses concerning these raw achievement scores were thus tested with data from each of the five Departmental Examinations.

Hypothesis 1. There are significant differences between the mean achievements of classes and the mean achievements of schools. This hypothesis was tested by forming the F - ratio MS_c/MS_r. This ratio was significant above the .01 level in all five subjects and the hypothesis was accepted.

Hypothesis 2. There are significant differences between the mean

achievements of schools and the mean achievements of districts. This hypothesis was tested by forming the F - ratio MS_s/MS_r. This ratio was significant above the .01 level in all five subjects, and the hypothesis was accepted.

Hypothesis 3. There are significant differences between the mean achievements of districts and the mean achievement of all pupils. This hypothesis was tested by forming the F - ratio MS_d/MS_r. This ratio was significant above the .01 level in all five subjects and the hypothesis was accepted.

Summary

This chapter has considered the findings of the analysis of variance of the raw scores of a sample of 2,069 pupils on five Alberta Departmental Examinations for Grade IX. This analysis permits total variation in achievement to be resolved into components representing sources from classes, schools, districts, and residuals. The residual source is formed from variations in pupil achievement within classes. Three between sources are formed from variations in classes, schools, and districts. All F - ratios in each of the five subjects were significant above the .01 level indicating that organizational variables in systems of education have demonstrable effects upon achievement.

In terms of variation, the largest source is the within variation from pupil achievement in classes. From sixty-three to sixty-nine per cent of total variation is associated with the residual source. The thirty-one to thirty-seven per cent of total variation associated with the three between sources is thus a measure of the influence of educational

systems upon achievement. The distribution of variation among all sources was similar in all five subjects. In general, the largest source of between variation was differences among classes and the smallest source differences among schools.

CHAPTER IX

FINDINGS RELATED TO THE ANALYSIS OF VARIANCE OF ACHIEVEMENT FACTOR SCORES

Factor scores are a measure of the amount of a factor possessed by an individual. When factors are orthogonal, the correlation of scores on a given factor with scores on any test in the battery is equal to the loading of the factor on the test. (Harman, p. 34). The factor scores thus represent the components of total test scores corresponding to factors defined in the analysis of the test battery.

In this study, two factors were identified and accepted as confirming hypotheses of the investigation. One of these was a large factor defined as a v:ed or achievement factor; the other was defined as a factor of socio-economic status. Two other factors were named but not accepted as confirming hypotheses of the study. Scores from the achievement factor were therefore available to test further hypotheses of the study. The findings of the analysis of variance of the raw achievement scores were reported in the previous chapter. This chapter reports the findings related to the analysis of variance of these achievement factor scores.

Achievement factor scores differ from raw achievement scores in that components due to error and specific factors have been removed from the total score. Moreover, components due to other common factors are also removed from the score. In this way, the achievement and socio-economic factor scores are independent of each other and of the remaining

common error, and specific factors. The achievement factor scores may also be regarded as a summary of the five achievement tests in the battery in that a single score instead of the scores from the five Departmental Examinations represents the achievement of an individual.

Factor Scores

Scores from four factors were derived from the matrix of rotated factors presented in Table VII. The matrix A given in this table contains seven factors. Only two of these were judged to be of significant size--the achievement and the socio-economic status factors. Two of the other named factors, had they been of significant size, would have been consistent with other hypotheses of the study. These factors were named according to the ability and study tests which had loadings upon them. Factor scores for all four factors were derived to demonstrate the determinacy of such scores using the method of image analysis. As will be shown in this chapter, even scores for very small factors conform closely to the theoretical form of such scores.

For convenience of reference, scores from these factors will be referred to as follows:

1. Factor I: v:ed

   This factor represents achievement and is named for the five Departmental Examinations and the three ability tests which have high loadings on it.

2. Factor II: s. e. s.

   This factor represents socio-economic status and is named for the three tests of socio-economic status which have moderate loadings on it.

3. Factor III: v - q

   This factor is named for the SCAT: Verbal and the SCAT: Quantitative tests which have equal and opposite loadings on it.

4. Factor IV: Study

   This factor is named for the test of study hours which has a single loading on it.

Only scores from the two significant factors were used in the analysis of variance.

The derivation of factor scores. The derivation of factor scores requires the inversion of the matrix A and the multiplication of A^{-1} W Z. The Matrix W is the matrix of regression weights for predicting scores on a test from the remaining n - 1 tests in the battery. The matrix Z is the matrix of standardized scores of 2,069 pupils on the thirteen tests in the battery.

The matrix A contains the rotated factor loadings of all thirteen factors extracted in the principal axes factoring of the image covariance matrix P. Table XV presents the inverse elements corresponding to the four factors mentioned above. Since scores on these factors were to be derived, only four rows of inverse elements were required from A^{-1}. These rows, corresponding to the four factors, are given in columnar form in Table XV. The pattern of weights to be used in the derivation of factor scores is observable in these elements.

In Factor I, the largest elements are found for the tests of achievement and educational ability with STEP: Reading and SCAT: Quantitative having the two largest elements. The tests of study and socio-economic status have negative elements.

TABLE XV

INVERSE ELEMENTS OF FOUR FACTORS OF THE MATRIX A

Variables	Factors			
	I	II	III	IV
Socio-economic Status	-.1660	1.0353	.2274	-.0668
Home Index	-.1646	.2739	.3829	-.0196
Occupation	-.1974	.8288	.0941	.0680
Study Hours	-.1674	-.0524	-.6154	3.1702
Motivation	.0858	-.0442	-.6391	-1.3354
STEP: Reading	.2014	.0227	-.1900	-.0174
Literature	.1275	-.0644	.4821	-.0412
Language	.1438	-.1856	.1594	-.0498
Social Studies	.1824	-.1065	.1033	-.0428
Mathematics	.1392	-.0179	-.4117	-.0640
Science	.1968	-.0285	-.8556	.0044
SCAT: Verbal	.1633	.0614	-2.1965	.0181
SCAT: Quantitative	.2139	-.0094	2.8783	-.0292

Factor II has large elements for the socio-economic tests and negative elements for the achievement examinations. Factor III has similar and opposite elements for the SCAT: Verbal and the SCAT: Quantitative tests. Factor IV has its single large element for the test of study hours.

Table XVI presents the regression weights of the Matrix W. This matrix contains the weights for predicting scores on a test from the n - 1 tests in the battery and is formed from:

$$W = I - S^2R^{-1}$$

as explained in Chapter VI.

The identity matrix I in the above formula eliminates each test sequentially from the rows of W. Thus a test is not used to predict itself and the diagonal elements of W are zero. For example, the weights for the prediction of Social Studies scores are given in the ninth row of W. The largest weight is .481 for the Science examination. This weight thus reflects the high correlation between scores on the Social Studies and Science examinations.

The matrix multiplication of A^{-1} by W produces the matrix of weights for predicting factor scores on each factor from the standardized scores of the thirteen variables. The multiplication of a row in A^{-1} by the thirteen column in W produces the thirteen test weights for the derivation of scores on a factor. Multiplication by successive rows of A^{-1} thus produces all the required test weights.

The test weights for four factors on each of the thirteen tests are presented in Table XVII. This matrix is of size 4 x 13 and is of interest because the test weights in a factor indicate the degree and direction of

TABLE XVI

THE MATRIX W

Variables[b]		Weights[a] 1	2	3	4	5	6	7	8	9	10	11	12	13
1.	S.E.S.	000	331	217	015	-006	-018	000	-047	016	002	013	067	014
2.	Home	273	000	350	006	029	061	054	-119	004	017	-025	177	000
3.	Occup.	201	393	000	017	-007	-042	058	110	-040	041	050	-012	-080
4.	Study	017	008	022	000	200	-114	295	116	-117	112	099	-214	-114
5.	Motiv.	-007	044	-010	217	000	-106	051	033	049	045	-057	-008	-033
6.	Rdg.	-007	030	-018	-040	-034	000	158	037	084	-34	-007	565	038
7.	Lit.	000	020	018	077	012	117	000	458	249	050	-001	114	-037
8.	Lang.	-016	-050	041	035	009	031	529	000	091	138	015	108	080
9.	Soc.St.	004	015	-012	-028	011	058	231	073	000	036	481	090	026
10.	Math.	000	006	013	030	011	026	051	121	039	000	453	-146	402
11.	Science	003	-008	015	023	-012	-005	000	-012	475	409	000	103	000
12.	V score	023	074	-004	-064	-002	479	131	107	111	-165	128	000	134
13.	Q score	006	000	-039	-045	-012	043	-056	105	042	599	000	178	000

[a]Decimal points omitted.

[b]See text for full names of variables.

TABLE XVII

THE MATRIX OF WEIGHTS FOR PREDICTING SCORES ON FOUR FACTORS

Variables	Factors			
	I	II	III	IV
Socio-economic Status	-.085	.244	.080	.072
Home Index	-.114	.677	-.081	-.027
Occupation	-.100	.311	.071	.060
Study Hours	.002	.007	-.104	-.294
Motivation	-.041	-.018	-.123	.632
STEP: Reading	.131	-.016	-.714	-.225
Literature	.101	-.066	-.579	.831
Language	.122	-.030	.110	.305
Social Studies	.216	-.050	-.384	-.456
Mathematics	.195	-.026	1.693	.267
Science.	.138	-.003	-.445	.346
SCAT: Verbal	.195	.093	.677	-.695
SCAT: Quantitative	.126	-.057	-.382	-.351

the contribution of tests to the factor score. The test weights for each factor are given in columnar form in Table XVII.

Factor I is the achievement factor. It has high positive loadings on the five achievement tests and the three academic ability tests indicating that high scores on these tests contribute positively to the factor score. The negative weights on the three tests of socio-economic status indicate that high scores on these tests contribute negatively to the achievement factor score. Conversely, the factor score of an individual who is low in socio-economic status is increased by these negative socio-economic weights. In this way, the achievement factor scores are independent of socio-economic status scores. That is, the correlation of the achievement factor scores have a theoretical zero correlation with socio-economic factor scores.

Similarly, scores on Factor II are influenced largely by the high positive weights associated with the three socio-economic tests. Weights on Factors III and IV indicate their dependence on complex relationships between the tests. This lack of simplicity makes the interpretation of scores on these factors uncertain.

The matrix Z was of size 13 x 2,069. That is, the matrix contains thirteen standardized test scores for each of the 2,069 pupils in the sample. Multiplication of the matrix in Table XVII by the matrix of standardized test scores produces the factor scores. Thus for a single factor, the row of test weights multiplied by the column of standard scores on thirteen tests produces the factor score of an individual. In this way, a factor score reflects not only the test weights but the performance of the individual on the tests. By the multiplication of

these matrices, sets of scores for each of the four factors are produced.

The theoretical distribution of each of these sets has a mean of zero and a standard deviation of unity. The means and standard deviations of these sets of factor scores are presented in Table XVIII. As may be seen in this table, these values closely approximate the theoretical parameters of factor scores. Thus with the method of image analysis, factor scores are determinate even for very small factors. The accuracy of the approximation to theoretical values of means and standard deviations is important if factor scores are to be uniquely determined. (Guttman, 1956, p. 276, p. 285). In conventional factor analysis, these values in factor scores are not always found with the result that such scores are not uniquely determined.

The Analysis of Variance of Factor Scores

For each of the factors described above, 2,069 factor scores were obtained. The achievement factor scores are relevant to hypotheses of this study. These scores were therefore subjected to analysis of variance in the same manner described for the raw achievement scores in the previous chapter. The socio-economic factor scores were also analyzed in this fashion although the analysis is not related to hypotheses of the study. No use was made of the scores on the other two factors other than to indicate the determinacy of scores from small factors.

Achievement factor scores. The results of the analysis of variance of the achievement factor scores are presented in Table XIX. The F - ratios are significant above the .01 level indicating system effects from classes, schools, and districts. The results are, however, similar

TABLE XVIII

MEANS AND STANDARD DEVIATIONS OF SCORES ON FOUR FACTORS

Factor	Mean	Standard deviation
I	.002	1.004
II	.005	.976
III	.001	1.002
IV	.001	1.002

TABLE XIX

ANALYSIS OF VARIANCE OF ACHIEVEMENT FACTOR SCORES

Source of variation	SS	Percentage of total variation	d of f	MS	F
Between classes	401.68	19.28%	44	9.129	12.8[a]
Between schools	58.39	2.80%	22	2.654	3.7[a]
Between districts	214.25	10.28%	21	10.202	14.3[a]
Residual	1,409.57	67.64%	1,981	.711	--
Total	2,083.90	100.00%	2,068	--	--

[a]Significant above .01.

to those obtained from the analysis of variance of the raw achievement scores. In terms of variation, approximately 68 per cent of the total sums of squares is associated with the residual pupil variation within classes; 32 per cent is associated with the total between sums of squares. Of the between variation, the pattern of its distribution among classes, schools, and districts is alsosimilar to that found in the analysis of the raw achievement scores. The largest proportion--19 per cent--is associated with class variation; 10 per cent is associated with district variation and the smallest proportion--3 per cent-- is associated with schools.

Socio-economic factor scores. The results of the analysis of variance of the achievement factor scores are presented in Table XX. The F - ratios are all significant above the .01 level. However, the F for classes approaches the value of 1.57 required for significance with these degrees of freedom. In view of possible departures from assumptions in these analyses of variance and since such departures indicate the increased probability of type - 1 errors, the significance of the ratio for classes is in doubt. However, these ratios and the percentages of variation in socio-economic factor scores are presented largely for comparison with the analysis of the achievement scores.

The distribution of variation among the five sources is quite different in the analyses of the achievement and socio-economic factor scores. In the analysis of the socio-economic factor scores, residual variation accounts for 77 per cent of the total sums of squares while only 23 per cent is associated with the total between variation. Of

TABLE XX

ANALYSIS OF VARIANCE OF SOCIO-ECONOMIC STATUS FACTOR SCORES

Source of variation	SS	Percentage of total variation	d of f	MS	F
Between classes	61.05	3.09%	44	1.387	1.8[a]
Between schools	78.56	3.98%	22	3.571	4.6[a]
Between districts	310.31	15.72%	21	14.777	19.2[a]
Residual	1,524.42	77.21%	1,981	.769	--
Total	1,974.36	100.00%	2,068	--	--

[a]Significant above .01.

this variation, the largest proportion is associated with districts--almost 16 per cent--with only relatively small proportions being associated with schools and classes. The largest portion of socio-economic differences in the sample is thus associated with district differences.

Discussion

The findings of the analysis of variance of the achievement factor scores confirm the findings of the analysis of variance of the raw achievement scores. In both cases, the systematic effects of classes, schools, and districts are established. The findings of these analyses thus confirm that educational systems are characterized by structure and organization consistent with the postulates of systems analysis. An educational system may thus be regarded as a hierarchy of sub-systems with each of these sub-systems having differential effects upon output. Had these effects not been established, the classes, schools, and districts would be regarded as largely meaningless units of organization. That is, these sub-systems would have to be regarded merely as a means of classification whose effects upon the outcomes of education is negligible or non-significant. Such sub-systems would therefore be irrelevant to the purposes of education.

Total achievement scores and factor achievement scores. The similarity of findings in the analysis of variance on the raw score achievement and that on the factor score achievement raises a question as to the reason for such similarity. The factor scores are not only purer measures of achievement in that error, specific factors, and other common

factors are partialed out; they are also independent of socio-economic effects. It might be presumed therefore that such scores would yield a smaller proportion of residual pupil variation than that associated with the raw score achievement. System effects would thus account for a larger proportion of total variation. The similarity of results found in the two types of analysis of variance does not support this presumption. The analysis of variance of achievement factor scores may be regarded as a summary of the analysis of raw achievement scores, despite some differences in patterns of variation found in the analyses of the different achievement tests.

Explanation of this similarity may be found in an examination of the factor matrix and in an examination of the matrix of weights for the derivation of factor scores.

The factor matrix, given in Table VII, indicates that the large achievement factor (Factor I) has substantial loadings on all the achievement and ability tests. These loadings are precisely the correlation between achievement factor scores and the total scores on these achievement and ability tests. The achievement factor scores are thus highly related to total scores on each of the achievement and ability tests. The difference between the achievement factor scores and raw scores of any achievement or ability test is apparently not great enough to alter the overall pattern of the distribution of total variation.

The matrix of weights for predicting factor scores from the thirteen tests in the battery is given in Table XVII. These weights reveal the same dependence of achievement factor scores upon the Departmental Examinations and the ability tests, even though such scores are independent of the socio-economic tests.

The importance of an independent measure of academic ability is thus indicated. The SCAT and STEP tests are apparently general achievement tests rather than independent measures of intellectual ability. The same abilities and knowledge which produce high achievement on the Departmental Examinations go to make high achievement on the SCAT and STEP Tests. SCAT is, of course, designed as a predictor of academic success but the claim is also made for it that it is not as specific as achievement tests nor as general as intelligence tests. (Educational Testing Service, Technical Manual). In the spectrum of ability tests, SCAT is presumed to stand mid-way between specific and general abilities. (Cronbach, p. 235). In relation to the Alberta Departmental Examinations at least, this separation is not evident.

Had the battery of tests used in this study included measures of independent intellectual ability, it may be presumed that the loadings of the Departmental Examinations, the SCAT tests, and the STEP test would have been lower with some of the variance now assigned to achievement being assigned to a general intellectual ability factor. The independence of the achievement factor from a factor of general intellectual ability was discussed in Chapter VII on the findings of the image analysis. Though it was argued there that v:ed factors are independent of general intellectual ability factors, this independence is not clearly established in this study. Had such independence been established, the achievement factor scores would have been less highly related to the Departmental Examinations and to the academic ability tests.

The size of F - ratios. The F - ratios reported in this chapter

and the previous chapter on the analysis of variance on raw score achievement are all of considerable size. The size of these ratios cannot be used to argue for the greater significance of one sub-system effect over another. That is, the ratios established only the significance of the various between sources of variance over the residual source. Differences in the various sources of variation were discussed in terms of the proportion of total sums of squares with which each was associated after the overall significance of the variances had been established.

The size of these F - ratios also gives confidence in regard to possible departures from the assumptions of analysis of variance. As has been indicated, only one ratio approached the minimum value for significance at the .01 level. This ratio involved the analysis of variance of the socio-economic factor scores and was not related to hypotheses of the study. All other F - ratios were substantially above the minimum requirements for significance at the .01 level. Since departures from the assumptions of analysis of variance have the effect of increasing the probability of type - 1 errors, the magnitude of the F - ratios give confidence that the statistical decisions made in the study do not involve such errors. That is, departures from the assumptions of analysis of variance tend to increase the size of the F - ratio needed to establish the significance of the ratio. There are indications that these increases are small even for cases where N is small. Where N is large, the effects of such departures are correspondingly smaller. (Linquist, pp. 78ff).

Thus the size of the obtained F - ratios and the size of the sample involved in this study give confidence to the statistical decisions made on the basis of these ratios.

Intraclass correlation. Difficulties in comparing F - ratios may be met by the use of the statistic R, known as the intraclass correlation. (Haggard). These difficulties arise from the dependence of F - values upon degrees of freedom and from the lack of a theoretical upper limit in such values. Thus F - ratios are comparable only in terms of their corresponding P - values unless the mean squares involved in each ratio have identical degrees of freedom. (Haggard, p. 6).

The F's reported in this study for the analysis of achievement data are all significant above the .01 level. It is thus impossible to discuss the relative importance of classes, schools, and districts as sources of variance in achievement in terms of the size of the F - ratios. Instead, these differences have been discussed in terms of their contributions to total variation. Intraclass correlations, however, make it possible to change F's into a statistic which permits the comparison of the contributions of between variances to total variances. R may be either positive or negative but has a maximum value of unity. Moreover, this statistic is independent of the number of degrees of freedom associated with the mean squares, though its size is related to the size of obtained F's.

> The coefficient of intraclass correlation can be defined as a ratio of two variances, the between-class variance to the total variance. . . .
>
> Maximal positive correlation exists when all the intraclass scores are identical and the scores differ only from class to class. As the relative heterogeneity of the intraclass scores increases, the computed values of R will decrease; maximal negative correlation exists when the heterogeneity of the intraclass scores is maximal and all the class means are the same. (Haggard, pp. 4,6).

In the analysis of the achievement factor scores, the intraclass correlations were calculated as .34 for classes, .10 for schools, and .36

for districts. The significance of thse results is established by the significance of their associated F - ratios. These correlations may be transformed into Fisher z coefficients and the significance of differences between them established. (Haggard, p. 14). Comparison of these correlations revealed no significant difference between classes and districts but both classes and districts were significantly greater sources of total variance than were schools.

Hypotheses

Three hypotheses of the study are related to the analysis of variance of the achievement factor scores.

Hypothesis 8: There are significant differences between the mean factor score achievement of classes and the mean factor score achievement of schools.

The mean square of between class variance was significantly greater than the mean square of residual variance within classes. This ratio was significant above the .01 level of confidence. The hypothesis was accepted.

Hypothesis 9: There are significant differences between the mean factor achievement of schools and the mean factor score achievement of districts.

The mean square of between school variance was significantly greater than the mean square of residual variance within classes. This ratio was significant above the .01 level of confidence. The hypothesis was accepted.

Hypothesis 10: There are significant differences between the mean factor score achievement of districts and the mean factor

score achievement of all pupils.

The mean square of between district variance was significantly greater than the mean square of residual variance within classes. This ratio was significant above the .01 level of confidence. The hypothesis was accepted.

Summary

This chapter has considered the findings of the analysis of variance on the achievement factor scores. The findings related to this analysis were similar to those resulting from the analysis of the raw achievement scores. All F - ratios were significant indicating effects associated with system differences in classes, schools, and districts. Approximately 68 per cent of total variation was associated with the residual variance of pupils within classes. Approximately 32 per cent was associated with the total between variation. Of the total between variation, the largest proportion was associated with class differences and the smallest with school differences. Classes and districts contributed larger proportions of between variance than did schools.

Scores on four factors were derived including two very small factors. The distributions of these scores were found to approximate closely their theoretical values indicating the determinateness of factor scores through image analysis even for very small factors. The analysis of the socio-economic factor scores indicated that the preponderance of socio-economic differences was associated with district variations.

The achievement factor scores were found to be closely related to raw achievement scores and to the tests of academic ability used in the battery of tests. The tests of academic ability used in the battery are

therefore not separable from the Departmental Examinations. Each group of tests seems to measure the same ability. In view of the test loadings on the achievement factor and of the weights determined for the derivation of achievement factor scores, it was concluded that the large achievement factor identified in the study likely contains variance which could have been assigned to a general intellectual factor had such a factor emerged in the analysis.

CHAPTER X

SUMMARY AND CONCLUSIONS

This study has dealt with two major problems. The first is suggested from the theory of systems analysis which regards organizations as complexities whose parts are related and directed towards some goal or output. Educational enterprises typically have an hierarchical structure composed of several sub-systems. Three common units of organization in educational systems are the classroom, the school, and the school district. It may be presumed from the postulates of systems analysis that each of these sub-systems has some relation to aims and purposes in education. One of the essential aims of education is to facilitate the acquisition of knowledge by pupils. An output related to this aim is the achievement of students on examinations. Thus one purpose of this study was to determine the contribution of three sub-systems in education to output as measured by pupil achievement on subject matter examinations.

This purpose suggests an analysis of variance design in which the effects of classes, schools, and districts are compared to a residual pupil variance within classes. A modified analysis of variance design was used to test these effects.

A second purpose of the study was to resolve pupil achievement into components associated with the sources of such achievement. These components would represent such variables in achievement as the intelligence, socio-economic status, motivation, and study hours of pupils.

A further important component would represent the impact of the educational system upon learning. The isolation of this component would permit its analysis in the same fashion as that proposed above for the total achievement scores of pupils on subject matter examinations.

Summary

This study involved the testing of a sample of 2,069 pupils on a battery of thirteen tests. The results on these tests were related through the theory of systems analysis to the effects of three educational sub-systems. Analysis of variance was used to assign the effects of these sub-systems. Image analysis and principal axes factoring were used to resolve pupil achievement into components and the component identified as achievement was related to these sub-systems using the same method of analysis.

Systems analysis. A system is defined as an organized complexity whose parts interact to produce some output or to achieve certain goals. Such systems differ from the typical systems in physical science which tend toward maximum disorder, entropy, and the degradation of energy. Biological or human systems have the characteristics of control, differentiation, purposiveness, and adaptiveness. They tend, so long as the entity is maintained, towards a homeostatic organization which is quite different from the entropy and disorganization which characterizes the end result of physical systems.

The consequences of open systems--the systems of human and biological organizations--are of considerable importance. A first consequence suggests that systems as "complexes of elements in interaction"

have general characteristics and laws of behavior which are quite different from the characteristics of any of these elements taken singly or in any combination apart from the total organization. This is the consequence of isomorphy and suggests the possibility of a general theory of organization. Such a general theory would attempt to explain why organizations behave as they do and why they often fail to perform in an expected or desired fashion. The theory would illuminate the mechanics of organization and permit the rational control of organizations. The tendency of bureaucratic organizations to assume a life and will of their own apart from their intended purposes is a consequence of the systematic control inherent in organizations. Knowledge of the mechanics of this control would thus permit its redirection towards the intended purposes of the organization.

Another consequence of open systems, related to the first, is the tendency of organizations towards equifinal and steady-state conditions. Equifinality is seen in the tendency of systems to produce a uniform output despite varying inputs. A steady state is the equilibrium achieved between a system and its environment. Systems exchange inputs and outputs with their environments and these inputs and outputs are modified by the steady state and equifinal conditions of the systems. Knowledge of what produces one type of equifinality or one steady state rather than a multitude of other conditions is thus important to any general theory of organization.

Control in systems is closely related to the flow of information within the system. Part of the output of open systems is fed back as an input. This feedback of output as information permits systems to be

self-regulating and self-directing. When the flow of information is obscured or obstructed, control becomes difficult. That is, the actual control exerted may be different from that which is intended. Knowledge of the flow of information and adequate feedback are thus essential to system control. Moreover, adequate information must be connected to adequate means for interpreting the information and for relating it to precise and predictable effectors.

A system may be regarded as a unit of analysis and is defined in terms of related elements which go to produce some outcome selected for its significance to a given purpose or goal. The use of systems theory made in this study involved the definition of an educational system composed of an hierarchical arrangement of classes, schools, and districts as sub-systems. The achievement of pupils on examinations of subject matter knowledge was selected as a significant output related to the aims of education.

The sample. The sample used in the study involved 2,069 pupils from twenty-two districts in the province of Alberta. Two schools from each district were selected yielding forty-four schools in the sample. Two classes of pupils in each school were also selected yielding eighty-eight classes in total. Objects within these sub-systems would include the financial resources of districts, the physical facilities of schools, organizational arrangements within classes, and the personnel associated with each of these sub-systems. Characteristics of these objects include the values of communities, attitudes and competencies of personnel, and the individual differences of pupils. Questionnaires sent to teachers

and principals revealed a wide variety of characteristics in each of these sub-systems. The distributions of pupil achievement on five Departmental Examinations revealed that the achievement of pupils in the sample closely approximated distributions in the province as a whole.

The tests. Data on a battery of thirteen tests were obtained for each of the 2,069 pupils in the sample. Approximately 150 pupils were dropped from the sample because of incomplete data. This number is made up largely of pupils who withdrew from school after the time the collection of data was begun. Other pupils not included in the total possible sample defined above were those absent from school at the time data were collected in a particular school and those for whom complete socio-economic data were not available.

The battery of tests was divided into four groups. Five tests, identified as achievement tests, were the Alberta Grade IX Departmental Examinations. These examinations are administered to all Grade IX students in Alberta in the final month of the school year. The results of these examinations are scaled and used as a basis for making educational and vocational decisions concerning each student. The achievement tests included examinations in Literature, Language, Social Studies, Mathematics, and Science.

A group of socio-economic tests were formed from three tests of social status. The Home Index is a scale of twenty questions concerning housing conditions, parental educational background, and the familial style of life. This questionnaire was completed by each pupil. The Occupation Index was formed from the reported occupations of parents of

the pupils in the sample. Using the Blishen Occupational Index, a score may be assigned to each occupation according to the years of training required for it and to the income associated with it. A third scale, called Socio-Economic Status was formed by the combination of equally weighted scores from the Home Index and Occupation scales.

A third group of tests was identified as academic ability. The group was made up of three tests. Two were sub-tests of SCAT. These sub-tests yield measures of quantitative and verbal ability. The third test in the group was the STEP: Reading which is a measure of reading ability.

The final group of tests, composed of two tests, was identified as educational motivation. One of these tests was a self-report by pupils on their hours of study. The other was a scale called Motivation which was derived from a questionnaire of sixteen items completed by each pupil. These sixteen items were subjected to scalogram analysis and the eleven best fitting items were selected to form the motivation scale. The reproducibility of this scale was below the arbitrary level of .90 but the distribution of item popularities and the absence of high order error patterns permitted its use as a quasi-scale.

Image analysis. Image analysis is a technique developed by Louis Guttman for the factoring of correlation matrices. It is superior to conventional factor analysis because it involves close approximations of communalities and yields a matrix for factoring which contains only the common parts of test variables. Factors derived in image analysis are concrete, real entities rather than hypothetical random variables.

Moreover, image analysis provides a method of extracting precise factor scores.

The correlation matrix of the thirteen tests was modified by the image analysis method and factored using the principal axes criterion. The principal axes factors were rotated to a simple structure position using the Varimax method. The factoring and rotation yielded thirteen factors only two of which were of significant size. The results of the factoring may be summarized as follows:

1. A large factor in the correlation matrix accounts for 72 per cent of common variance and 40 per cent of total variance in the thirteen tests. This factor had high loadings on five Departmental Examinations of subject matter knowledge. Lower, but still high, loadings on this factor were also found for the three tests of academic ability. The factor was identified as a v:ed or achievement factor.

2. The only other factor of significant size in the correlation matrix accounted for 15 per cent of common variance and 8 per cent of total variance. The three tests of socio-economic status had high loadings on this factor. The factor was identified as socio-economic status.

3. Two other factors hypothesized in the study did not appear as a result of the factoring of the correlation matrix. These were an intellectual ability factor and a factor of educational motivation. Three factors related to these hypotheses were identified but their small size precluded the acceptance of hypotheses on this basis.

The analysis of variance of raw achievement scores. Scores from

the five Grade IX Departmental Examinations were subjected to the analysis of variance. This analysis permits the total variation in scores to be assigned to sources from classes, schools, districts, and residual variation within classes. The analysis of the five sets of examination scores in this fashion yielded similar results in each case, although some differences among the different subjects were apparent. F - ratios were formed by comparing the between sources of variance in classes, schools, and districts with the residual variance within classes. All F - ratios in each of the five analyses were significant above the .01 level of confidence. The substantial size of these ratios and the large number of pupils in the sample gave confidence that statistical decisions on the basis of these ratios did not involve type-1 errors, despite the possible violation of the assumptions of analysis of variance. In each of these analyses the largest source of variation was the residual source within classes. From 63 to 69 per cent of variation was associated with this source. The three between sources thus accounted for from 31 to 37 per cent of total variation. In general, the largest between source of variation was classes with districts representing the next largest. By far the smallest source was schools. The findings of the five analyses of variance on raw score achievement may be summarized as follows:

1. Classes are a significant source of variance in the achievement of pupils in educational systems as measured by examinations in academic subjects.

2. Schools are a significant source of variance in the achievement of pupils in educational systems as measured by examinations in academic subjects.

3. Districts are a significant source of variance in the achievement of pupils in educational systems as measured by examinations in academic subjects.

The derivation of factor scores. Scores for four factors identified in the image analysis of the correlation matrix were derived. Two sets of these scores were derived from the factors identified as achievement and socio-economic status. These were the two significant factors resulting from the image analysis. Scores were also derived for two small factors. The weights for the derivation of these scores from the thirteen tests corroborated the interpretation placed on the large significant factors. In these cases, the weights associated with tests had a simple pattern. In the achievement factor, the achievement and ability tests had large positive weights while the socio-economic tests had moderate negative weights. On the socio-economic factor the largest weights were observed for the three socio-economic tests while all other weights were near zero. In the case of the two small factors, the patterns of test weights were complex indicating the difficulty of placing any interpretation upon scores from such factors. Each set of scores from the four factors conformed closely to their theoretical mean and standard deviation values indicating the determinateness of factor scores through image analysis even for very small factors.

The analysis of variance of achievement factor scores. The scores derived from the large achievement factor were subjected to analysis of variance in the same manner as that described for the raw achievement scores. The factor score of achievement may be regarded as a component of the total achievement score. This total score is composed of various sources in the individual differences of pupils and error. The analysis of the achievement factor score only thus partials out of the total score

other common factors, error, and specific factors in tests. It might be presumed that the analysis of such a score would reduce residual pupil variation in relation to between variation thus leaving larger system effects associated with classes, schools, and districts.

However, the results of the analysis of variance of the achievement factor scores did not produce results that were strikingly different from the analysis of the raw achievement scores. Of the total variation, 68 per cent was associated with residual variation within classes; 32 per cent was associated with the total between variation. Of the total between variation, 19 per cent was associated with class differences, 10 per cent with district differences, and 3 per cent with school differences. These percentages most closely resemble the distribution of variation in the analysis of the raw Science achievement scores.

F - ratios in this analysis were formed in the same manner as that described for the analysis of the raw achievement scores. Each of these three ratios was significant above the .01 level of confidence. The ratios were again substantially above those required for significance at this level.

The findings of the analysis of variance of the achievement factor scores may be summarized as follows:

1. Classes are a significant source of variance in the achievement of pupils in educational systems as measured by an achievement factor score.

2. Schools are a significant source of variance in the achievement of pupils in educational systems as measured by an achievement factor score.

3. Districts are a significant source of variance in the achievement of pupils in educational systems as measured by an achievement factor score.

4. Classes and districts have significantly higher intraclass correlations than schools.

Conclusions Related to Image Analysis and Principal Axes Factoring

The conclusions related to image analysis and principal axes factoring have to do largely with techniques to be applied in the isolation of an achievement factor as a criterion for judging the effects of educational systems.

1. Image analysis was revealed as a superior method for the factoring of correlation matrices. This method, though it has been known for some years, has been little used in studies appropriate to the factoring of matrices. The method is computationally easy and requires little additional information than is usually calculated in conventional factor analyses. The determinate factors and factor scores of image analysis recommend it for much wider application than it has received.

2. The factoring of matrices involving subject matter examinations and standardized tests of educational ability and achievement yields a large factor on which such tests have high loadings. The determination of such a factor permits the use of the factor scores associated with it as a criterion for assessing the effects of educational systems. Many studies in educational research now use some form of a pupil gain score to assess the effectiveness of teachers, the impact of administrative policies, and the results of curriculum change. Factor

scores on achievement are therefore suggested as a ready measure of output in education.

3. Pupil achievement as measured by standardized achievement tests or school examinations contains a significant component due to a socio-economic factor. The socio-economic factor identified in this study accounted for 15 per cent of common variance and 8 per cent of total variance. In dealing with pupils in rural areas, the occupation index of social status fails to make distinctions among pupils whose parents are farmers. The need for more tests that are independent predictors of socio-economic status is also indicated if the communalities of these tests are to be raised.

4. The achievement factor scores derived in this study were highly related to the total achievement scores of pupils on subject matter examinations. It is therefore possible that some of the variance in the achievement factor could be assigned to an independent factor of general intellectual ability. Such a factor could be established by inclusion in the test battery of measures saturated with Spearman's g. Several tests measuring this factor are available. The combination of such tests with subject matter examinations and measures of socio-economic status would permit the isolation of a factor relatively independent of the total achievement score but perhaps reflecting the impact due to educational systems more closely.

5. The measures of educational motivation included in this study failed to define a factor related to them. This failure is probably due to unreliability and error in the tests. It may also be due to the independence of motivation and educational achievement. In this case

motivation must be viewed as being composed of specific factors largely unrelated to achievement and socio-economic tests. As in the case of the socio-economic tests, the communalities of the motivation tests could be raised by the inclusion of independent measures of motivation. It may be concluded, too, that there are other attitudes and values in pupils besides motivation which may be more highly related to achievement.

6. The SCAT and STEP tests are closely related to the Grade IX Departmental Examinations even though the SCAT tests were designed to be less specific than educational achievement tests. The SCAT tests are likely to be good predictors of achievement in the Alberta school system, though their potential as measures of general intellectual ability is placed in doubt.

Conclusions Related to Scalogram Analysis

1. The items selected for scalogram analysis in this study failed to define a consistent attitude towards education. The quasi-scale which was defined was not validated by large correlations with any other variable in the study. Only low correlations were found with the tests of study hours and the achievement tests. The causes of these low correlations are uncertain although error and unreliability are undoubtedly part of the source.

2. The items used to form the scale apparently reflect various aspects of attitude towards education. More specific items centering on activities related to achievement might form a more consistent scale. The conclusions of the Miller study also indicate the multi-dimensionality of educational motivation.

Conclusions Related to the Analysis of Variance of Achievement Scores

The conclusions related to the analysis of variance of the achievement scores are perhaps best presented in relation to the null hypotheses of the analysis of variance. These null hypotheses may be stated as follows:

$$\mu_c = \mu_s = \mu_d = \mu$$

where c refers to classes, and s to schools and d to districts. That is, the null hypothesis is that the population means of classes, schools, and districts are equivalent to the grand mean of all pupil achievement. The only source of between variance would thus be error and all F - ratios would be non-significant. Class, school, and district variances would thus be equally good estimates of the general population variance.

If this hypothesis were established, the educational system defined in this study would represent a degenerate system in which energy was equally dispersed with entropic forces at a maximum. The classification of pupils into classes, schools, and districts would be a meaningless organization since none of these sub-systems would be related to output. Moreover, the establishment of the null hypotheses would mean that residual variance within classes would be at a maximum and the between variance at a minimum. Therefore whether a pupil was educated in one class, school, and district or in quite different ones would have no effect on his achievement. That is, differences in pupil achievement would serve to form within class variance but would have little effect upon variance between classes, schools, and districts.

None of the null hypotheses of the analysis of variance were confirmed in either the analysis of raw score achievement or the analysis

of the achievement factor scores. The estimates of between variance differed among classes, schools, and districts. But in each case these were significant sources of variance.

As a proportion of the total variation, the residual source from pupils within classes is large--up to 69 per cent in the case of the analysis of Science achievement. However, the 31 per cent of variation associated with the three between sources in this case is still significant. Patterns of distribution of variation differed among the different achievement scores but in general the results were similar.

Conclusions from these analyses may be stated as follows:

1. Classes are a significant source of variance in pupil achievement. Classes therefore form a sub-system within the general educational system. Objects within this sub-system, the characteristics of these objects, and the relationships between them have effects upon the output of the educational system as a whole.

2. Schools are a significant source of variance in pupil achievement. Schools, too, form a sub-system and influence the output of the total system.

3. Districts are a significant source of variance in pupil achievement. Districts are thus a third significant sub-system within educational systems.

4. Classes and districts are significantly larger sources of achievement variance in school systems than are schools.

5. The analysis of achievement factor scores may be used as a general indicator of the effectiveness of school systems. In this study the analysis of achievement factor scores assigned 32 per cent of total

variation to between sources in classes, schools, and districts. Between variation in classes accounted for 19 per cent of the total variation; districts accounted for 10 per cent of total variation; and schools accounted for 3 per cent of total variation.

6. The variance assigned to classes, schools, and districts may include variance due to the intellectual abilities of pupils within them. An achievement score which can be shown to be independent of intellectual, socio-economic, and attitude factors in pupils would be a theoretically desirable criterion for the evaluation of educational system efficiency.

7. The achievement factor scores used in this study as a criterion of system effectiveness are independent of the socio-economic dimension in achievement. They are also independent of other common, error, and specific factors in achievement tests. This independence recommends them as a more reliable criterion than total achievement scores from examinations or standardized tests.

Implications of the Study

The findings and conclusions of this study have several implications for educational research. Perhaps the most important implication of the study is the tentative map which it offers as a guide to the operation of educational systems. The nature of systems theory involves the analysis of complex wholes. Systems theory thus provides a means of observing the action of large educational units. Much research in education has been research in small and the need to examine effects of total operations is indicated if a theory of educational organization is to be developed. This study provides an indication of where the largest

effects of educational systems are to be found. For example much research has been devoted to the evaluation of teacher effectiveness. It is apparent from this study that classroom effects are only part of the forces which go to produce achievement in pupils. Some account must therefore be taken of the complex interrelations which have a bearing upon classroom achievement if the effectiveness of a teacher is to be judged by the performance of pupils in a class. Moreover, the teacher is only one of many objects within the classroom sub-system which can have an influence upon the achievement of pupils. Therefore, even if classroom sources of variation can be isolated from other sources, there remains the problem of which variables in classes are significant in relation to achievement.

The study indicates, too, that classes and districts have the largest influence of the three sub-systems considered in the study upon pupil achievement. A further task for administrative theory is thus to identify significant objects in various levels of the system. The characteristics of these objects must be specified and related to the outcomes of educational systems.

Systems analysis is important, too, in that it is meaningless without the definition of aims and purposes in a system. Moreover, it is essential in systems analysis to specify a desired output and to relate this output to the aims of the system and to its method of operation. It is impossible to decide whether systems or policies and practices within them are "good" unless a criterion of evaluation is established. Much knowledge in the administration of education is traditional knowledge or the knowledge based on classification systems. A theory of organization

demands much more precise knowledge of the operation of systems than that available from subjective evaluation of organizational operation. Systems analysis answers the question of how organizations do work rather than the question of how they ought to work. This study has attempted to evaluate an educational system in terms of a significant output from it; its findings indicate that there are systemic effects associated with output. The question of how these effects are produced is still unclear. This problem is therefore suggested as an area for further research.

Another implication of this study concerns the flow of information in an educational system and the role of feedback in the control of the system. As was indicated in the discussion of systems theory, the flow of information and the maintenance of feedback is essential to the operation of systems. One of the most vital kinds of information needed in education concerns the effects of systems on output. This kind of information is frequently at a minimum in educational systems since few means are provided for the measurement of effects. If it can be known that educational systems affect educational outputs in certain ways and to certain degrees, it is then possible to ask if these effects are the desired ones and to ask whether the effects are large enough. Knowledge of this kind, if it is given to the administrator (the controller of operations in systems), enables modifications of the system to be made to increase or to adjust the effects of a system. If certain means are not adequate to produce desired effects, then the need to develop different or improved means is indicated. It is also possible to ask what amounts and kinds of information are sent to the control mechanism in the system, and to ask how well this mechanism is able to read and interpret

the information provided. In this way, it is possible to evaluate the effectiveness of the administrator himself.

In general, then, systems analysis is a means to establish a theory of organization; in particular it can provide knowledge needed to establish a theory of educational administration. This study offers some findings relevant to the gross effects of educational systems. It indicates, as well, some techniques which may be used in further studies applying the systems analysis model.

Further Research

The need for the following types of investigation are suggested from this study.

1. The analysis of variance model used in this study is well adapted to problems involving the assignment of organizational effects on achievement. The model should be used again with other degrees of freedom. The total number of students involved may possibly be decreased as may be the number of districts involved. The number of schools per district and classes per school should be increased.

2. An attempt should be made to compare educational organizations having matched or controlled inputs and environmental conditions. In this way, a stringent test of equifinal and steady state conditions may be made.

3. Studies which attempt to account for variability at various organizational levels are necessary. Outputs of organizations with desirable characteristics should be compared with those having less desirable characteristics.

4. Significant objects in educational organization and their characteristics should be identified. That is, the organizational variables having the greatestinfluence on achievement should be isolated.

5. The establishment of a criterion of educational efficiency remains a basic, unresolved problem. Image analysis of achievement with independent measures of intelligence and the derivation of factor scores as a measure of efficiency is indicated as a possible criterion.

BIBLIOGRAPHY

BIBLIOGRAPHY

Ackerman, Walter I. "Teacher Competence and Pupil Change," Harvard Educational Review, 24:273-89, Fall, 1954.

Anderson, C. C. "The Evaluation of Teaching Proficiency: Research Findings and Theoretical Considerations." Unpublished paper read to the Canadian Association of Professors of Education, June, 1962. (Mimeographed.)

_____ and S. M. Hunka, "Teacher Evaluation: Some Problems and a Proposal," Harvard Educational Review, 33:74-95, Winter, 1963.

Baird, Richard. "The Politics of Education: Practice and Theory: Controversy and Control," The Canadian Administrator, 1:31-38, April-May, 1962.

Bantock, G. H. "Educational Research: A Criticism," Harvard Educational Review, 31:265-280, Summer, 1961.

Barber, Bernard. Social Stratification. New York: Harcourt, Brace, and Company, 1957.

Barr, A. S. et al. The Measurement of Teaching Ability. Madison: Dembar Publications, 1945.

Bidwell, Charles E. "A New Dilemma in Administration," Harvard Educational Review, 26:388-404, Fall, 1956.

Black, D. B. "The Prediction of University Freshman Success Using Grade IX Departmental Scores," Alberta Journal of Educational Research, 4:229-239, December, 1959.

Blishen, Bernard R. "Construction and Use of an Occupational Scale," Canadian Journal of Economics and Political Science, 24:519-531, November, 1958.

Bloom, Benjamin S. (ed.). Taxonomy of Educational Objectives. New York: David McKay, 1956.

Boulding, Kenneth. "General Systems Theory--The Skeleton of Science," General Systems: Yearbook for the Advancement of General Systems Theory, 1:11-17, 1956.

_____. The Organizational Revolution. New York: Harpers, 1953.

Brehaut, Willard. "Educational Research in Canada, 1956-1958," Canadian Education, 15:37-40, June, 1960.

British Columbia. Report of the Royal Commission on Education. Victoria: The Queen's Printer, 1960.

Buros, Oscar Krisen, (ed.). The Fifth Mental Measurement Yearbook. New Jersey: The Gryphon Press, 1959.

Bushnell, Don D. and John F. Cogswell. A Computer-Based Laboratory for Automation in School Systems. Santa Monica: Systems Development Corporation, 1961.

Canada. Bureau of Statistics. Classification of Occupations. Ottawa: Queen's Printer, 1951.

Canadian Education Association. "Summaries of Graduate Theses in Education, 1959-1961," Canadian Education and Research Digest, June, 1960; 1961; 1962.

Cattell, Raymond B. "Three Basic Factor-Analytic Research Designs," Psychological Bulletin, 49:499-520, September, 1952.

Coleman, James S. The Adolescent Society. Glencoe: The Free Press, 1961.

Cooley, William C. "Research Methodology and Modern Data Processing," Harvard Educational Review, 31:257-263, Summer, 1961.

Cooperative Study of Secondary School Standards. Evaluative Criteria. Washington: Cooperative Study Committee, 1950.

Cronbach, Lee J. Essentials of Psychological Testing. New York: Harper and Brothers, 1960.

Dailey, John T. "The Role of Feedback in a Structured Social-Servo System," Paper delivered to the Behavioral Sciences Conference, University of New Mexico, July, 1957. (Mimeographed.)

_____. "A System for Classifying Public High Schools," Paper delivered to AERA and AASA annual meeting, Atlantic City, New Jersey, February, 1962. (Mimeographed.)

Davis, Allison. Social-Class Influences Upon Learning. Cambridge: Harvard University Press, 1957.

Davis, Frederick B. "The Interpretation of Principal-Axis Factors," Journal of Educational Psychology, 38:471-478, December, 1947.

Downey, Lawrence W. The Task of Public Education. Chicago: Midwest Administration Center, University of Chicago, 1960.

_____, and Frederick Enns (eds.). The Social Sciences and Administration. Edmonton: University of Alberta, 1963.

Dubin, Robert. Human Relations in Administration. Englewood Cliffs: Prentice-Hall, 1961.

Dunlop, G. M., S. Hunka and H. Zingle. "Individual Differences in Alberta Schools," Alberta Journal of Educational Research, 4:5-14, December, 1955.

Educational Testing Service. Cooperative School and College Ability Tests: Manual for Interpreting Scores. Princeton: Educational Testing Service, 1957.

_____. Cooperative School and College Ability Tests: Technical Report. Princeton: Educational Testing Service, 1957.

Edwards, Allen Lewis. Experimental Design in Psychological Research. New York: Rinehart, 1960.

Elley, Warwick Bartram. "A Comparative Analysis of the Socio-Economic Bias in Selected Intelligence Tests." Unpublished Ph.D. thesis, University of Alberta, 1961.

Enns, Frederick (ed.). "The Aims of Education in Canada." A report compiled for the Canadian Conference on Education, 1962. (Mimeographed.)

Eyman, R. K., H. F. Dingman, and C. E. Meyers. "Comparison of Some Computer Techniques for Factor Analytic Rotation," Educational and Psychological Measurement, 22:201-214, Spring, 1962.

Flanagan, John C. "Data Processing in Large-Scale Research Projects," Harvard Educational Review, 31:250-256, Summer, 1961.

Gordon, Milton M. Social Class in American Sociology. Durham: Duke University Press, 1958.

Gordon, Wayne C. The Social System of the High School. Glencoe: The Free Press, 1957.

Gough, H. "A Short Social Status Inventory," Journal of Educational Psychology, 40:52-56, 1949.

Green, Bert F. "A Method of Scalogram Analysis Using Summary Statistics," Psychometrika, 21:79-88, 1956.

Greer, Scott A. Social Organization. New York: Random House, 1955.

Griffiths, Daniel E. Administrative Theory. New York: Appleton-Century-Crofts, 1959.

Guilford, J. P. Psychometric Methods. New York: McGraw-Hill Book Company, Inc., 1954.

_____. "When Not To Factor Analyze," Psychological Bulletin, 49:26-37, January, 1952.

Guttman, Louis, "'Best Possible' Systematic Estimates of Communalities," Psychometrika, 21:273-285, September, 1956.

_____. "Image Theory for the Structure of Quantitative Variates," Psychometrika, 18:277-296, December, 1953.

_____. "Multiple Rectilinear Prediction and the Resolution into Components," Psychometrika, 5:75-99, March, 1940.

_____. "A New Approach to Factor Analysis: The Radex," Mathematical Thinking in the Social Sciences, Paul F. Lazarsfeld, (ed.). Glencoe, Illinois: The Free Press, 1954.

_____. "Structural Theory for Intergroup Beliefs and Action," American Sociological Review, 24:318-328, June, 1959.

_____. "What Lies Ahead in Factor Analysis?" Educational and Psychological Measurement, 18:497-515, Autumn, 1958.

Haggard, Ernest A. Intraclass Correlations and the Analysis of Variance. New York: Dryden Press, 1958.

Hall, A. D., and R. E. Fagen. "Definition of a System," General Systems: Yearbook for the Advancement of General Systems Theory, 1:18-28, 1956.

Halpin, Andrew W. (ed.). Administrative Theory in Education. Chicago: Midwest Administration Center, 1958.

_____. "A Paradigm for Research on Administrator Behavior," Administrative Behavior in Education, Roald F. Campbell and Russell T. Gregg, editors. New York: Harper and Brothers, 1957.

Harman, Harry H. Modern Factor Analysis. Chicago: University of Chicago Press, 1960.

Hearn, Gordon. Theory Building in Social Work. Toronto: University of Toronto Press, 1958.

Herriott, Robert E. "Some Social Determinants of Educational Aspiration," Harvard Educational Review, 33:157-177, Spring, 1963.

Hukins, Austin Adolphus. "A Factorial Investigation of Measures of Achievement of Objectives in Science Teaching." Unpublished Ph.D thesis, University of Alberta, 1963.

Jenkinson, Marion D. "Some Implications of Psychological Research for Elementary Education," Canadian Education and Research Digest, 2: 195-205, September, 1962.

Kahl, J. A., and J. A. Davis, "Comparison of Indices of Socio-Economic Status," American Sociological Review, 20:317-325, 1956.

Kaiser, Henry F. "The Best Approximation of a Common-Factor Space." Bureau of Educational Research, University of Illinois. (Mimeographed.)

Kershaw, J. A., and R. N. McKean. Systems Analysis and Education. Santa Monica: The Rand Corporation, 1959.

Lazarsfeld, Paul F., and Morris Rosenberg, (eds.). The Language of Social Research. Glencoe, Illinois: The Free Press, 1955.

_____, and Daniel E. Griffiths. "The Social Sciences and the Study of Administration: A Rationale," The Social Sciences and Administration. Lawrence W. Downey and F. Enns, (eds.). Edmonton: Division of Educational Administration, University of Alberta, 1963.

Leeson, Daniel N., and Donald L. Dimitry. Basic Programming Concepts and the IBM 1620 Computer. New York: Holt, Rinehart and Winston, 1962.

Linquist, Everet Franklin. Design and Analysis of Experiments in Psychology and Education. Boston: Houghton Mifflin, 1953.

Lord, Frederic M. Inferring True Scores from Fallible Scores. Princeton: Educational Testing Service, 1960.

MacArthur, R. S., and S. Hunka. School Examination Practices and Standards in Alberta. Edmonton: Alberta Advisory Committee on Educational Research, University of Alberta, 1958.

_____, and S. A. Linstedt. The Alberta Teacher Force in 1957-58. Edmonton: The Alberta Advisory Committee on Educational Research, University of Alberta, 1958.

Maccia, George S. An Educational Theory Model: General Systems Theory. Columbus, Ohio: Center for the Construction of Theory in Education, 1962.

Mauch, J. "A Systems Analysis Approach to Education," Phi Delta Kappan, 43:158-162, January, 1962.

Mayer, Martin. The Schools. New York: Harper and Brothers, 1961.

Mead, George Herbert. The Social Psychology of George Herbert Mead. Anselm Strauss, (ed.). Chicago: University of Chicago Press, 1959.

Miklos, Erwin. "Dimensions of Conflicting Expectations and the Leader Behavior of Principals." Unpublished Doctoral thesis, University of Alberta, 1963.

Miller, James G. "Towards a General Theory for the Behavioral Sciences," The American Psychologist, 10:513-531, 1955.

Mitchell, James V. Jr. "An Analysis of the Factorial Dimensions of the Achievement Motivation Concept," Journal of Educational Psychology, 52:179-187, August, 1961.

Mitzel, Harold E. A Behavioral Approach to the Assessment of Teacher Effectiveness. New York: Office of Research and Evaluation, Division of Teacher Education, College of the City of New York, 1957.

_____, and Cecily F. Gross, "A Critical Review of Pupil Growth Criteria in Studies of Teacher Effectiveness," Division of Teacher Education, Board of Higher Education of the City of New York, New York, 1956. (Mimeographed.)

Nash, Paul. "Objective Examinations and the Process of Education," Canadian Education and Research Digest, 2:99-104, June, 1962.

Presthus, Robert. The Organizational Society. New York: Alfred A. Knopf, 1962.

Project Talent. "A National Inventory of Aptitudes and Abilities," Washington: Project Talent office of the University of Pittsburgh, Bulletin No. 1, November, 1959, Bulletin No. 2, June, 1960, Bulletin No. 3, December, 1960.

Reiss, Albert J., and Albert L. Rhodes. A Social-Psychological Study of Adolescent Conformity and Deviation. Washington: U.S. Department of Health, Education and Welfare, U.S. Office of Education.

Scheffé, Henry. The Analysis of Variance. New York: John Wiley and Sons, 1959.

Stouffer, Samuel A., et al. Measurement and Prediction. Vol. IV of Studies in Social Psychology in World War II. 4 vols. Princeton: Princeton University Press, 1950.

Swanson, Austin D. Effective Administrative Strategy. New York: Teachers College, Columbia University, 1961.

Tate, Merle W., and Charlotte E. Voss. "A Study of the Davis-Eells Test of Intelligence," Harvard Educational Review, 26:374-387, Fall, 1956.

Thomas, J. Alan. "Efficiency in Education: An Empirical Study," Administrator's Notebook, 11:No. 2, October, 1962.

Thurstone, Louis Leon. Multiple-Factor Analysis. Chicago: University of Chicago Press, 1947.

_____. The Nature of Intelligence. London: Kegan Paul, 1924.

Vernon, P. E. Intelligence and Attainment Tests. London: University of London Press, 1960.

_____. The Measurement of Abilities. London: University of London Press, 1956.

_____. The Structure of Human Abilities. London: Methuen and Co., 1950.

von Bertalanffy, Ludwig. "General System Theory," General Systems: Yearbook for the Advancement of General Systems Theory, 1:1-10, 1956.

_____, and Anatol Rapoport, (eds.). General Systems: Yearbook of the Society for the Advancement of General Systems Theory. Vol. I, 1956.

Walker, Helen M., and Joseph Lev. Statistical Inference. New York: Holt, Rinehart and Winston, 1953.

Walton, John. "The Theoretical Study of Educational Administration," Harvard Educational Review, 25:169-178, Summer, 1955.

Warner, W. Lloyd. Social Class in America: A Manual for the Measurement of Social Status. New York: Harper and Brothers, 1960.

Wiener, Norbert. The Human Use of Human Beings. New York: Doubleday, 1954.

Winer, B. J. Statistical Principles in Experimental Design. New York: McGraw-Hill Book Company, 1962.

APPENDIX A

THE ANALYSIS OF VARIANCE

I. THE ANALYSIS OF VARIANCE[a]

Let X_{ijkL} be the mark of the L^{th} pupil in the k^{th} class of the j^{th} school in the i^{th} district, ($i = 1,\ldots, 22$; $j = 1, 2$; $k = 1, 2$; $L = 1,\ldots, N_{ijk}$),

where N_{ijk} = the number of pupils in the k^{th} class of the j^{th} school in the i^{th} district.

Let $N_{ij} = \sum_k N_{ijk}$ = the number of pupils in the j^{th} school of the i^{th} district,

$N_i = \sum_j N_{ij}$ = the number of students in the i^{th} district,

$N = \sum_i N_i$ = the total number of pupils in the entire sample.

Let $\bar{X} = 1/N \sum_i \sum_j \sum_k \sum_L X_{ijkL}$ = the average mark for all pupils in the sample,

$\bar{X}_i = 1/N_i \sum_j \sum_k \sum_L X_{ijkL}$ = the average mark for pupils in the i^{th} district,

$\bar{X}_{ij} = 1/N_{ij} \sum_k \sum_L X_{ijkL}$ = the average mark for pupils in the j^{th} school in the i^{th} district,

$\bar{X}_{ijk} = 1/N_{ijk} \sum_L X_{ijkL}$ = the average mark of pupils in the k^{th} class of the j^{th} school in the i^{th} district.

---

[a]The author is indebted to Dr. J. R. McGregor of the Department of Mathematics, University of Alberta for the derivation of this model of the analysis of variance.

The total sum of squares about the grand mean is

$$S_t = \sum_i \sum_j \sum_k \sum_L (X_{ijkL} - \bar{X})^2.$$

S_t may be partitioned into

$$S_t = S_r + S_c + S_s + S_d,$$

where

$$S_r = \sum_i \sum_j \sum_k \sum_L (X_{ijkL} - \bar{X}_{ijk})^2,$$

$$S_c = \sum_i \sum_j \sum_k N_{ijk} (\bar{X}_{ijk} - \bar{X}_{ij})^2,$$

$$S_s = \sum_i \sum_j N_{ij} (\bar{X}_{ij} - \bar{X}_i)^2,$$

$$S_d = \sum_i N_i (\bar{X}_i - \bar{X})^2,$$

S_r = the residual SS,

S_c = SS between classes within schools and districts,

S_s = SS between schools within districts,

S_d = SS between districts.

The analysis of variance then has the form:

Source of variation	SS	d of f	MS
Between classes	S_c	$IJ (K - 1)$	$S_c/IJ (K - 1)$
Between schools	S_s	$I (J - 1)$	$S_s/I (J - 1)$
Between districts	S_d	$I - 1$	$S_d/ (I - 1)$
Residual	S_r	$N - IJK$	$S_r/ (N - IJK)$
Total	S_t	$N - 1$	--

II. COMPUTATION FORMULAS[a]

The notation used in this section is identical to that used in the previous section with the following additions:

Let $T_{ijk} = \sum_{L} X_{ijkL}$ = the sum of achievement scores in the k^{th} class of the j^{th} school in the i^{th} district,

$T_{ij} = \sum_{k}\sum_{L} X_{ijkL}$ = the sum of achievement scores in the j^{th} school in the i^{th} district,

$T_{i} = \sum_{j}\sum_{k}\sum_{L} X_{ijkL}$ = the sum of achievement scores in the i^{th} district,

$T = \sum_{i}\sum_{j}\sum_{k}\sum_{L} X_{ijkL}$ = the sum of achievement scores in the entire sample.

X^2_{ijkL} = the square of the achievement score for the L^{th} pupil in the k^{th} class in the j^{th} school of the i^{th} district.

From these terms, values for the following expressions may be calculated:

(1) $\sum_{i}\sum_{j}\sum_{k}\sum_{L} X^2_{ijkL}$,

(2) $\sum_{i}\sum_{j}\sum_{k} (T_{ijk})^2/N_{ijk}$,

(3) $\sum_{i}\sum_{j} (T_{ij})^2/N_{ij}$,

(4) $\sum_{i} (T_{i})^2/N_{i}$,

(5) $(T)^2/N$.

---

[a]The author is indebted to Mrs. P. Montgomery of the University of Alberta Computing Center for the derivation of these formulas.

Sums of the squares for the sources of variation specified in the model are then calculated as follows:

$$SS_r = \sum_i \sum_j \sum_k \sum_L X^2 - \sum_i \sum_j \sum_k (T_{ijk})^2 / N_{ijk},$$

$$SS_c = \sum_i \sum_j \sum_k (T_{ijk})^2 / N_{ijk} - \sum_i \sum_j (T_{ij})^2 / N_{ij},$$

$$SS_s = \sum_i \sum_j (T_{ij})^2 / N_{ij} - \sum_i (T_i)^2 / N_i,$$

$$SS_d = \sum_i (T_i)^2 - (T)^2 / N$$

Values for the five required expressions may be calculated by performing successive one-way analyses of variance over classes, schools, and districts.

APPENDIX B

QUESTIONNAIRES AND RESPONSES

A Systems Analysis of Structure and Organization in Education

Principal's Questionnaire

Every effort has been made to hold the information given in this questionnaire in strict confidence.

It is not necessary to place your name on the questionnaire. School and district names are required only for classification of data.

The analysis of this information will be done using I.B.M. cards and data processing equipment. After responses have been coded and transferred to I.B.M. cards, complete anonymity is assured.
Please reply to each item.

A. IDENTIFICATION OF PARTICIPATING CLASSES

1. It is necessary to select two Grade IX classes from this school for participation in this study. The pupils in these classes and their teachers in Social Studies, Mathematics, English, and Science will also fill out questionnaires. IDENTIFY these two participating classes BY NUMBER OR LETTER.

   1. Class number ..... Class enrolment .....

   2. Class number ..... Class enrolment .....

2. Name of School .....

   .....

3. Name of District .....

   .....

B. SCHOOL CHARACTERISTICS

4. In what TYPE OF SCHOOL are you principal?

   CHECK ONE for appropriate type or closest type given.

   1 .6. Junior High School (Grades VII to IX)

   2 .10. Elementary and Junior High School

   3 .12. Junior High and Senior High

   4 .16. Elementary, Junior High, and Senior High

5. In what TYPE OF ADMINISTRATIVE UNIT is your school?

   CHECK ONE

   1 .32. In a school division or country

   2 .10. In a city district

   3 .2. In a town district

   4 ..... In a separate school district

   5 ..... In a village district

   6 ..... In a rural school district

6. What is the present ENROLMENT of pupils in the school where you are principal?

   CHECK ONE

   1 .2. 1 to 249

   2 .5. 250 to 349

   3 .11. 350 to 449

   4 .7. 450 to 549

   5 .13. 550 to 649

   6 .2. 650 to 749

   7 .3. 750 to 849

   8 ..... 850 to 999

   9 ..... 1000 or more

7. Does your school have a GYMNASIUM or GYMNASIUM-AUDITORIUM?

   CHECK ONE

   1 .44. Yes

   2 ..... No

8. How many TEACHERS are there in your SCHOOL? Include yourself as principal in the total.

CHECK ONE

1 1 1 to 6 teachers
2 ____ 7 to 9 teachers
3 2 10 to 12 teachers
4 6 13 to 15 teachers
5 7 16 to 19 teachers
6 10 20 to 24 teachers
7 16 25 to 34 teachers
8 2 35 to 49 teachers
9 ____ 50 teachers or more

9. What is the exact PUPIL-TEACHER RATIO in this school? Include yourself as principal in the total of teachers.

WRITE IN RATIO (Number of pupils divided by number of teachers and carry to one place of decimals.)

---

10. How OLD is your school BUILDING? (Give age of largest section of the school if parts are of different ages.)

CHECK ONE

1 3 1 to 2 years
2 17 3 to 5 years
3 10 6 to 8 years
4 8 9 to 12 years
5 2 13 to 16 years
6 ____ 17 to 20 years
7 ____ 21 to 25 years
8 ____ 26 to 29 years
9 4 30 years or more

C. SCHOOL ORGANIZATION

11. Are the Grade IX pupils GROUPED either by ABILITY or by ACHIEVEMENT?

CHECK ONE

1 19 Yes
2 25 No

12. Do the Grade IX TEACHERS hold STAFF MEETINGS BY GRADE?

CHECK ONE

1 15 Yes
2 29 No

13. Are there designated DEPARTMENT HEADS is your school for academic subjects such as English and Mathematics?

CHECK ONE

1 2 Yes
2 42 No

14. Does your school have a CENTRAL SCHOOL LIBRARY?

CHECK ONE

1 33 Yes
2 11 No

15. (If "yes" in No. 14, complete the following two questions.) Approximately how many BOOKS are in the SCHOOL LIBRARY?

CHECK ONE

1 ____ 100 or less
2 ____ 101 to 200 books
3 ____ 201 to 300 books
4 3 301 to 500 books
5 3 501 to 750 books
6 6 751 to 1000 books
7 9 1001 to 2000 books
8 12 2001 books or more

16. Does your school have a LIBRARIAN?

CHECK ONE

1 18 Yes
2 23 No

17. Does your school have TEAMS playing regularly in INTER-SCHOOL SPORTS?

CHECK ONE

1 37 Yes
2 7 No

D. ADMINISTRATIVE CHARACTERISTICS

18. Are you designated as a TEACHING PRINCIPAL? (Answer "yes" if any part of your time is regularly spent in teaching.)

CHECK ONE

1 37 Yes
2 6 No

19. How many DAYS PER WEEK of your time are given to ADMINISTRATION or SUPERVISION in your school? (Answer only for the time school is in session.)

CHECK ONE

1 4 One-half day or less
2 2 One day
3 2 One and one-half days
4 8 Two days
5 11 Two and one-half days
6 4 Three days
7 3 Three and one-half days
8 4 Four and one-half days
9 6 Five days

20. Do you receive SECRETARIAL HELP in your office?

CHECK ONE

1 29 Yes
2 15 No

21. HOW MUCH SECRETARIAL HELP do you receive per week?

CHECK ONE

1 16 One-half day or less
2 ___ One day
3 4 One and one-half days
4 4 Two days
5 3 Two and one-half days
6 ___ Three days
7 ___ Three and one-half days
8 1 Four and one-half days
9 15 Five days

E. TEACHING AND ADMINISTRATIVE EXPERIENCE

22. Counting the present school year, what is the TOTAL YEARS OF EXPERIENCE you have had as TEACHER and PRINCIPAL?

CHECK ONE

1 ___ 4 years or less
2 2 5 to 8 years
3 7 9 to 12 years
4 4 13 to 16 years
5 5 17 to 20 years
6 4 21 to 23 years
7 8 24 to 27 years
8 7 28 to 31 years
9 7 32 years or more

23. What is the TOTAL YEARS OF EXPERIENCE you have had as PRINCIPAL?

CHECK ONE

1 2 1 year
2 2 2 to 3 years
3 8 4 to 6 years
4 7 7 to 9 years
5 7 10 to 12 years
6 7 13 to 15 years
7 5 16 to 18 years
8 3 19 to 20 years
9 3 21 years or more

24. How long have you been PRINCIPAL IN THIS SCHOOL?

CHECK ONE

1 8 1 year
2 2 2 years
3 5 3 or 4 years
4 15 5 to 7 years
5 8 8 to 10 years
6 3 11 to 13 years
7 3 14 years or more

F. YOUR QUALIFICATIONS

25. What GENERAL ALBERTA TEACHING CERTIFICATE do you hold? Check one the MOST RECENT CERTIFICATE issued.

CHECK ONE

1 ___ Junior E.
2 ___ Standard E.
3 ___ Junior E. and I.
4 5 First class
5 ___ Second class
6 ___ Letter of authority
7 ___ Standard S.
8 32 Professional Certificate
9 7 Other (Please write in)

26. What is the EXTENT of your TOTAL ACADEMIC and PROFESSIONAL PREPARATION BEYOND HIGH SCHOOL?

CHECK ONE

1 ___ Less than a standard 1-year program (7 months) in a Normal School, Teachers' College, or University
2 1 A standard 1-year program (7 months or more) in a Normal School, Teachers' College, or University
3 ___ 2 complete years but less than 3 in a University and/or Teachers' College
4 ___ 3 complete years but less than 4 in a University and/or Teachers' College
5 12 4 complete years but less than 5 in a University and/or Teachers' College
6 13 5 complete years but less than 6 in a University and/or Teachers' College
7 18 6 or more complete years in a University and/or Teachers' College

27. What is the HIGHEST UNIVERSITY DEGREE you hold?

CHECK ONE

1 1 No degree
2 2 B. A.
3 1 B. Sc.
4 17 B. Ed.
5 10 Two or more Bachelor degrees
6 1 M. A.
7 10 M. Ed.
8 1 M. Sc.
9 1 Other (Please write in)

28. Have you taken GRADUATE COURSES in educational administration at College or University?

CHECK ONE

1 29 Yes
2 15 No

29. Do you hold a University degree in EDUCATIONAL ADMINISTRATION?

CHECK ONE

1 5 Yes
2 39 No

30. Have you ever attended the ALBERTA LEADERSHIP COURSE for school principals?

CHECK ONE

1 27 Yes
2 17 No

G. PERSONAL CHARACTERISTICS

31. Please indicate your sex.

1 44 Male
2 ____ Female

32. What is your age (Last birthday)?

1 ____ 25 years or under
2 ____ 26 to 30 years
3 6 31 to 35 years
4 4 36 to 40 years
5 5 41 to 45 years
6 17 46 to 50 years
7 8 51 to 55 years
8 3 56 to 60 years
9 1 61 years or more

33. What will be the total amount of your SCHOOL SALARY as principal for the school year 1961-1962 (September 1 to August 31)? Report the gross amount prior to withholding for taxes, retirement, or other items. Include any allowances for administrative duties, etc.

CHECK ONE

1 ____ $3,000 or less
2 ____ $3,001 to $3,500
3 ____ $3,501 to $4,000
4 ____ $4,001 to $5,000
5 1 $5,001 to $6,000
6 ____ $6,001 to $7,000
7 1 $7,001 to $8,000
8 6 $8,001 to $9,000
9 36 over $9,000

Your co-operation in completing this questionnaire is mueh appreciated.

Completeness in answering all items is important. Would you please check to see that a response has been made to all items.

You are reminded that all answers given here are held in strict confidence and that after coding, all replies beeome completely anonymous and are used thercafter only for statistical analysis.

THANK YOU.

COMMENTS (if any) _____

A Systems Analysis of Structure and Organization in Education

Teacher's Questionnaire

Every effort has been made to hold the information given in this questionnaire in strict confidence.

It is not necessary to place your name on the questionnaire. School and district names are required only for the classification of data.

The analysis of this information will be done using I.B.M. cards and data processing equipment. After responses have been coded and transferred to I.B.M. cards, complete anonymity is assured.

A. CLASS IDENTIFICATION

Two GRADE IX CLASSES in your school have been selected for participation in this research study.

THESE CLASSES ARE

1. Class ..... (identify by number or letter of the class)
2. Class ..... (identify by number or letter of the class)
3. CHECK THE COURSES below which you teach to these classes.

Class ..... English .....

Social Studies .....

Mathematics .....

Science .....

Class ..... English .....

Social Studies .....

Mathematics .....

Science .....

B. ALBERTA TEACHER CERTIFICATION

4. What GENERAL ALBERTA TEACHING CERTIFICATE do you now hold? Check only the LATEST CERTIFICATE ISSUED.

CHECK ONE

1 14 Junior E.

2 6 Standard E.

3 17 Junior E. and I.

4 11 First Class

5 ..... Second Class

6 4 Letter of Authority

7 25 Standard S.

8 68 Professional Certificate

9 ..... Other (please write in)

C. FORMAL ACADEMIC AND PROFESSIONAL PREPARATION

5. What is the extent of your TOTAL ACADEMIC AND PROFESSIONAL PREPARATION BEYOND HIGH SCHOOL?

CHECK ONE

1 2 Less than a standard 1-year program (7 months) in a Normal School, Teachers' College, or University

2 33 A standard 1-year program (7 months or more) in a Normal School, Teachers' College, or University

3 25 2 complete years but less than 3 in a University and/or Teachers' College

4 16 3 complete years but less than 4 in a University and/or Teachers' College

5 42 4 complete years but less than 5 in a University and/or Teachers' College

6 21 5 complete years but less than 6 in a University and/or Teachers' College

7 7 6 or more complete years in a University and/or Teachers' College

6. What is the HIGHEST UNIVERSITY DEGREE you hold?

CHECK ONE

1 75 No degree

2 12 B.A.

3 4 B.Sc.

4 37 B.Ed.

5 13 Two or more bachelor degrees

6 1 M.A.

7 1 M.Sc.

8 3 M.Ed.

9 ..... Other (please write in)

D. UNIVERSITY OR COLLEGE CONTENT COURSES

7. How many University or College CONTENT COURSES have you completed in HISTORY, ECONOMICS, AND GEOGRAPHY? (Two half-courses should be reported as one full course.)

CHECK ONE

1 26 None

2 23 One course

3 32 Two courses

4 22 Three courses

5 16 Four courses

6 12 Five courses

7 4 Six courses

8 1 Seven courses

9 6 Eight courses or more

8. How many University or College CONTENT COURSES have you completed in ENGLISH? (Two half-courses should be reported as one full course.)

CHECK ONE

1 25 None

2 58 One course

3 26 Two courses

4 14 Three courses

5 12 Four courses

6 4 Five courses

7 1 Six courses

8 ..... Seven courses

9 1 Eight courses or more

9. How many University or College CONTENT COURSES have you completed in SCIENCE i.e. physical and biological sciences? (Two half-courses should be reported as one full course.)

CHECK ONE

1 37 None

2 34 One course

3 33 Two courses

4 12 Three courses

5 3 Four courses

6 3 Five courses

7 3 Six courses

8 3 Seven courses

9 9 Eight courses or more

10. How many University or College CONTENT COURSES have you completed in MATHEMATICS? (Two half-courses should be reported as one full course.)

CHECK ONE

1 61 None

2 29 One course

3 13 Two courses

4 14 Three courses

5 12 Four courses

6 1 Five courses

7 2 Six courses

8 ..... Seven courses

9 1 Eight courses or more

E. METHODS COURSES

11. Have you taken METHODS COURSES in the teaching of SOCIAL STUDIES? Count only credit courses taken at University, College, or Normal School.

CHECK ONE

1 95 Yes

2 42 No

12. Have you taken METHODS COURSES in the teaching of SCIENCE? Count only credit courses taken at University, College, or Normal School.

CHECK ONE

1 81 Yes

2 49 No

13. Have you taken METHODS COURSES in the teaching of MATHEMATICS? Count only credit courses taken at University, College, or Normal School.

CHECK ONE

1 87 Yes

2 44 No

14. Have you taken METHODS COURSES in the teaching of ENGLISH? Count only credit courses taken at University, College, or Normal School.

CHECK ONE

1 87 Yes

2 49 No

F. YOUR TEACHING EXPERIENCE

15. Counting the present school year, what is the TOTAL NUMBER OF SCHOOL YEARS of full-time teaching experience you have had?

CHECK ONE

1 12 1 year

2 9 2 years

3 14 3 to 4 years

4 15 5 to 7 years

5 21 8 to 11 years

6 29 12 to 16 years

7 17 17 to 21 years

8 19 22 to 30 years

9 8 31 years or more

16. Counting the present year, how many YEARS have you taught IN THIS SCHOOL?

CHECK ONE

1 36 1 year beginning in September

2 2 1 year beginning after September

(Write in month you began .....)

3 28 2 years

4 7 3 years

5 18 4 to 5 years

6 23 6 to 8 years

7 16 9 to 12 years

8 8 13 to 17 years

9 6 18 years or more

17. How many COMPLETE YEARS of academic and/or professional training beyond high school did you have BEFORE YOU STARTED TEACHING?

CHECK ONE

1 16 less than one full year

2 77 1 year

3 13 2 years

4 9 3 years

5 22 4 years

6 7 5 years

7 2 6 years or more

G. YOUR TEACHING PREFERENCES

18. In which of the following SUBJECT AREAS of the Grade IX curriculum would you PREFER TO TEACH?

CHECK ONE

1 24 English

2 32 Social Studies

3 23 Science

4 29 Mathematics

19. In which of the following SUBJECT AREAS of the Grade IX curriculum do you consider that you are MOST ADEQUATELY PREPARED TO TEACH? Count any kind of formal preparation which you consider useful.

CHECK ONE

1 27 English

2 34 Social Studies

3 22 Science

4 24 Mathematics

H. PERSONAL CHARACTERISTICS

20. Please indicate your sex.

CHECK appropriate answer.

1 99 Male

2 46 Female

21. What is your age (Last birthday)?

CHECK ONE

1 22 25 years or under

2 12 26 to 30 years

3 28 31 to 35 years

4 18 36 to 40 years

5 21 41 to 45 years

6 26 46 to 50 years

7 11 51 to 55 years

8 5 56 to 60 years

9 2 61 years or more

22. What will be the total amount of your SCHOOL SALARY as a teacher for the school year 1961-1962 (September 1 to August 31)? Report the gross amount prior to withholding for taxes, retirement, etc.

CHECK ONE

1 1 $3,000 or less

2 6 $3,001 to $3,500

3 8 $3,501 to $4,000

4 37 $4,001 to $5,000

5 27 $5,001 to $6,000

6 25 $6,001 to $7,000

7 19 $7,001 to $8,000

8 16 $8,001 to $9,000

9 6 Over $9,000

Your co-operation in completing this questionnaire is much appreciated.

Completeness in answering all items is important. Would you please check to see that a response has been made to all items.

You are reminded that all answers given here are held in strict confidence and that after coding, all replies become completely anonymous and are used thereafter only for statistical analysis.

THANK YOU.

COMMENTS (if any) .....

THE PUPIL ATTITUDE QUESTIONNAIRE

The following questions ask you to give an opinion about school and school life.

Read each statement and decide whether you agree or disagree with it.

EXAMPLE: The work we do in school is too hard.

Agree___ ___ ___ ____ ___Disagree
+2 +1 0 -1 -2

Show how strong your opinion is by checking ANY ONE of the blanks between AGREE and DISAGREE. If you check a +2, you show a strong agreement with the statement. If you check a -2, you show a strong disagreement with the statement.

Checking the zero shows a neutral feeling.

CHECK ONE for each statement.[a]

1. All students should go to school for as many years as possible.
2. Even though school is hard, it will help me to be a better person in later life.
3. A school is just as important in a community as a hospital.
4. Most teachers try to help students as much as they can.
5. Getting a good education is worth more than making a lot of money.
6. Some students don't appreciate the opportunities they have in school.
7. High school is not enough education to be successful in life.
8. School should not be compulsory after Grade 8.

---

[a]Each statement on the original questionnaire had a five-point scale beneath it as in the example statement given above.

9. Working hard in school is more important than having a good time.
10. Going to school is not the only way to learn important things.
11. Getting good marks in school is more important than being popular.
12. A good athlete should be allowed to spend more effort on sports than on school work.
13. A student who gets good marks in school is not usually popular.
14. A boy who refuses to try out for school teams is showing poor school spirit.
15. Girls should support school teams by attending home games.
16. Doing your homework is more important than going out to have a good time.

TABLE XXI (A)

DISTRIBUTION OF PUPIL RESPONSES TO ITEMS OF A SCALE OF ATTITUDE TO SCHOOLING

Item	Direction of scaled items and percentages of respondents endorsing items strongly
1. All students should go to school for as many years as possible.	Positive (64%)
Agree 1,337 (+2) 409 (+1) 180 (0) 105 (-1) 46 (-2) Disagree	
2. Even though school is hard, it will help me to be a better person later in life.	Positive (76%)
Agree 1,580 (+2) 389 (+1) 65 (0) 28 (-1) 18 (-2) Disagree	
3. A school is just as important in a community as a hospital.	Positive (54%)
Agree 1,130 (+2) 560 (+1) 240 (0) 109 (-1) 40 (-2) Disagree	
4. Most teachers try to help students as much as they can.	Positive (42%)
Agree 867 (+2) 822 (+1) 231 (0) 104 (-1) 54 (-2) Disagree	
5. Getting a good education is worth more than making a lot of money.	Positive (36%)
Agree 749 (+2) 620 (+1) 496 (0) 118 (-1) 97 (-2) Disagree	
6. Some students don't appreciate the opportunities they have in school.	Positive (48%)
Agree 1,039 (+2) 708 (+1) 218 (0) 79 (-1) 37 (-2) Disagree	

TABLE XXI (b)

DISTRIBUTION OF PUPIL RESPONSES TO ITEMS
OF A SCALE OF ATTITUDE TO SCHOOLING

Item	Direction of scaled items and percentages of respondents endorsing items strongly
7. High school is not enough education to be successful in life. Agree 373 (+2) 570 (+1) 464 (0) 381 (-1) 291 (-2) Disagree	Positive (18%)
8. School should not be compulsory after Grade 8. Agree 133 (+2) 158 (+1) 334 (0) 410 (-1) 1,030 (-2) Disagree	Negative (49%)
9. Working hard in school is more important than having a good time. Agree 582 (+2) 704 (+1) 497 (0) 209 (-1) 85 (-2) Disagree	Positive (28%)
10. Going to school is not the only way to learn important things. Agree 743 (+2) 768 (+1) 303 (0) 172 (-1) 91 (-2) Disagree	Not scalable
11. Getting good marks in school is more important than being popular. Agree 520 (+2) 673 (+1) 581 (0) 227 (-1) 89 (-2) Disagree	Positive (25%)

TABLE XXI (c)

DISTRIBUTION OF PUPIL RESPONSES TO ITEMS
OF A SCALE OF ATTITUDE TO SCHOOLING

Item	Direction of scaled items and percentages of respondents endorsing items strongly
12. A good athlete should be allowed to spend more effort on sports than on school work.	Not scalable
Agree 103 (+2) 208 (+1) 389 (0) 752 (-1) 627 (-2) Disagree	
13. A student who gets good marks in school is not usually popular.	Not scalable
Agree 219 (+2) 328 (+1) 391 (0) 579 (-1) 565 (-2) Disagree	
14. A boy who refuses to try out for school teams is showing poor school spirit.	Not scalable
Agree 376 (+2) 461 (+1) 413 (0) 459 (-1) 360 (-2) Disagree	
15. Girls should support school teams by attending home games.	Not scalable
Agree 705 (+2) 776 (+1) 456 (0) 93 (-1) 36 (-2) Disagree	
16. Doing your homework is more important than going out to have a good time.	Positive (36%)
Agree 744 (+2) 793 (+1) 352 (0) 127 (-1) 52 (-2) Disagree	

TABLE XXII

DISTRIBUTION OF SCORES ON THE ELEVEN ITEM QUASI-SCALE OF EDUCATIONAL MOTIVATION

Score	Frequency (N = 2,069)
11	17
10	47
9	105
8	141
7	211
6	279
5	309
4	296
3	236
2	228
1	119
0	81
Mean	4.78
S. D.	2.50

TABLE XXIII

DISTRIBUTION OF STUDY HOURS
AS REPORTED BY PUPILS

How much time, on the average, do you spend on homework outside of school hours?

CHECK ONE

1	108	None or almost none
2	405	About $\frac{1}{2}$ hour per day
3	719	About 1 hour per day
4	595	About $1\frac{1}{2}$ hours per day
5	242	About 2 hours or more per day

THE HOME INDEX QUESTIONNAIRE

1. What is your father's occupation?

_____

(Be clear. For example: A sales clerk at Eaton's, door to door salesman for Fuller Brush, travelling salesman for Massey-Ferguson.)

2. Mother's occupation _____

In the following questions, mark your answer by putting a circle in the right place. For example, in the question "Does your family own a car?" draw a circle around the Yes if your family has a car, and around the No if it does not. Be sure to answer all questions.

1. Does your family own a car? . . . . . Yes No
2. Does your family have a garage or carport? . . . . . Yes No
3. Did your father go to high school? . . . . . Yes No
4. Did your mother go to high school? . . . . . Yes No
5. Did your father go to university? . . . . . Yes No
6. Did your mother go to university? . . . . . Yes No
7. Is there a writing desk in your home? . . . . . Yes No
8. Does your family have a hi-fi or record player? . . . . Yes No
9. Does your family have a piano? . . . . . Yes No
10. Does your family get a daily newspaper . . . . . Yes No
11. Do you have your own room at home? . . . . . Yes No
12. Does your family own its own home? . . . . . Yes No
13. Is there an encyclopedia in your home? . . . . . Yes No
14. Does your family have more than 100 hard-covered books? (e.g. 4 shelves 3 feet long) . . . . . Yes No
15. Did your parents borrow any books from the library in the last year? . . . . . Yes No

16. Does your family leave town each year for a holiday?. . Yes No

17. Do you belong to any club where you have to pay fees? . Yes No

18. Does your mother belong to any clubs or organizations such as study, church, art, or social clubs? . . . . . Yes No

19. Does your father belong to any such clubs or organizations? . . . . . Yes No

20. Have you ever had lessons in music, dancing, art, swimming, etc., outside of school? . . . . . Yes No

APPENDIX C

DISTRIBUTIONS OF ACHIEVEMENT TEST SCORES

TABLE XXIV

DISTRIBUTIONS OF PUPIL ACHIEVEMENT
IN FIVE GRADE IX ACADEMIC SUBJECTS

	Subjects (N = 2,069)				
Interval	English Literature	English Language	Social Studies	Mathematics	Science
95 - 99		11	5	16	10
90 - 94		44	54	48	44
85 - 89		99	87	114	96
80 - 84		148	116	141	138
75 - 79		157	176	158	181
70 - 74		195	194	203	189
65 - 69		213	172	218	173
60 - 64		183	137	166	182
55 - 59		148	152	171	180
50 - 54	1	177	198	187	159
45 - 49	37	320	264	298	272
40 - 44	204	209	331	214	279
35 - 39	337	85	112	47	105
30 - 34	364	41	51	30	40
25 - 29	347	20	18	31	14
20 - 24	582	9	1	10	5
15 - 19	130	10	1	10	2
10 - 14	58			6	
5 - 9	8			1	
0 - 4	1				
Mean	28.92	60.21	58.86	60.62	59.82
S. D.	8.15	16.27	16.22	16.54	16.17

APPENDIX D

DISTRIBUTIONS OF SOCIO-ECONOMIC TEST SCORES

TABLE XXV

DISTRIBUTIONS OF SCORES ON THE OCCUPATION SCALE AND THE SOCIO-ECONOMIC STATUS SCALE

Interval	Occupation scale	Interval	Socio-Economic Status Scale (N = 2,069)
160 - 169	1		
150 - 159	12		
140 - 149	14	80.0 - 84.9	13
130 - 139	48	75.0 - 79.9	10
120 - 129	152	70.0 - 74.9	5
110 - 119	222	65.0 - 69.9	33
100 - 109	472	60.0 - 64.9	138
90 - 99	684	55.0 - 59.9	221
80 - 89	370	50.0 - 54.9	149
70 - 79	82	45.0 - 49.9	1122
60 - 69	12	40.0 - 44.9	354
50 - 59		35.0 - 39.9	24
40 - 49			
30 - 39			
Mean	100.9		50.10
S. D.	27.9		6.66

TABLE XXVI

DISTRIBUTION OF SCORES ON THE HOME INDEX QUESTIONNAIRE

Score	Frequency (N = 2,069)
20	8
19	27
18	48
17	80
16	93
15	133
14	144
13	187
12	199
11	201
10	208
9	193
8	187
7	124
6	108
5	73
4	31
3	15
2	10
Mean	11.02
S. D.	3.65

APPENDIX E

DISTRIBUTIONS OF EDUCATIONAL ABILITY TEST SCORES

TABLE XXVII

DISTRIBUTIONS OF SCORES ON THREE TESTS
OF EDUCATIONAL ABILITY

Interval	STEP: Reading	SCAT: Verbal	SCAT: Quantitative (N = 2,069
60 - 64		6	
55 - 59		128	
50 - 54	1	239	10
45 - 49	22	287	250
40 - 44	266	316	421
35 - 39	377	354	405
30 - 34	392	271	416
25 - 29	574	218	283
20 - 24	229	145	171
15 - 19	158	88	77
10 - 14	45	17	30
5 - 9	4		6
0 - 4			
Mean	29.51	38.42	34.54
S. D.	8.29	10.82	8.59

B29814